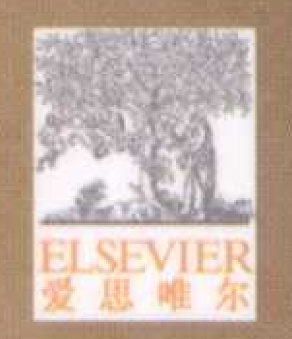
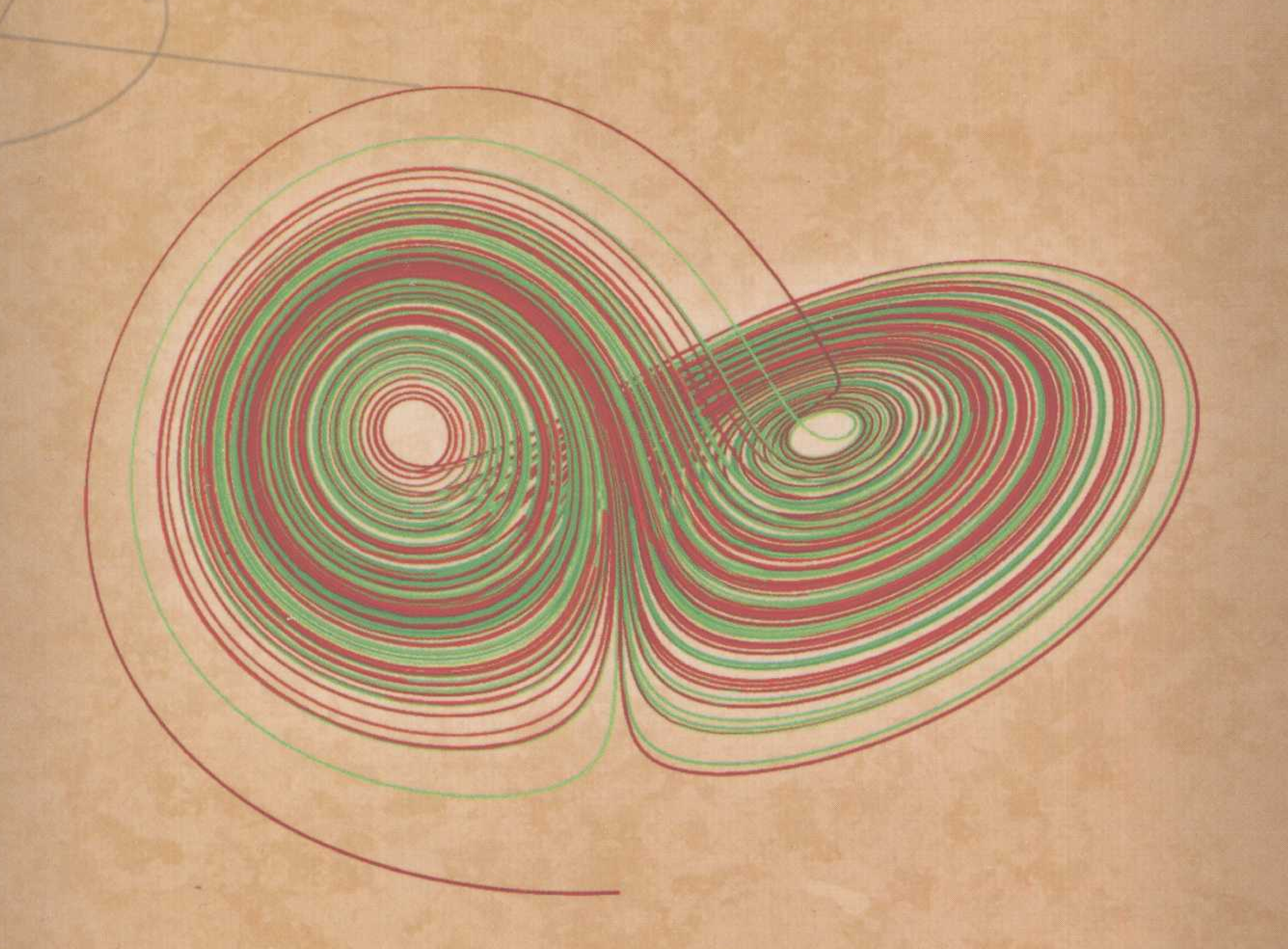
TURING 图灵数学·统计学丛书 19







Differential Equations, Dynamical Systems, and an Introduction to Chaos

微分方程、动力系统与混沌导论

（第 2 版）

［美］ Morris W. Hirsch
Stephen Smale 著
Robert L. Devaney
甘少波 译



微分方程、动力系统与混沌导论（第2版）

Differential Equations, Dynamical Systems, and an Introduction to Chaos

"数学著作的典范。……熟悉第1版的读者会发现新版涵盖面更宽，适用面更广。"

——美国数学月刊

"本书出自3位世界最杰出的动力系统专家之手，他们阐述数学原理的能力超强。书中涵盖了目前许多研究都要用到的混沌基准模型。"

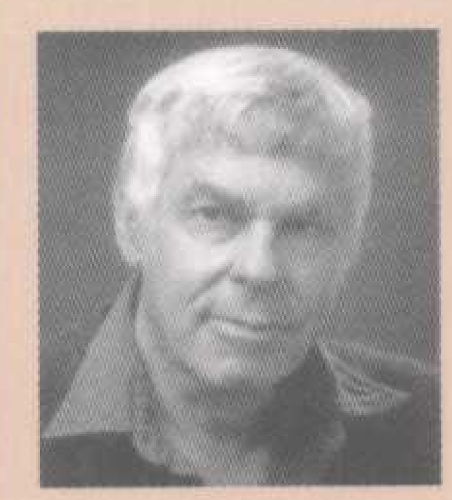
—— Bruce Peckham，明尼苏达大学

本书是经典的微分方程和动力系统著作，由3位世界知名数学家合著而成。第1版于30多年前出版，在科学界产生了深远的影响。新版反映了这一数学领域的巨大变化，论述了各种新的课题及其在化学、电子工程、天体力学和生态学等诸多学科中的广泛应用和影响。

新版修订时充分考虑了各学科的需求，适当降低了对数学背景的要求，使本书适用面更加广泛。不仅可供应用数学和其他理工科专业师生用作微分方程后续课程的教材，也是帮助科技工作者解决实际工作中遇到的微分方程课题的重要参考书。



Morris William Hirsch 加州大学伯克利分校荣休教授，世界著名的动力系统专家，单调动力系统理论创始人。他是1982年菲尔兹奖得主William Paul Thurston的博士导师。除本书外，他还著有*Differential Topology*一书。



Stephen Smale 20世纪最杰出的数学家之一，因解决了广义庞家莱猜想于1966年荣获菲尔兹奖，并于2007年荣获沃尔夫奖。美国科学院、艺术和科学院院士，1996年获美国国家科学奖。他在微分拓扑、动力系统、非线性分析、数理经济学等领域都做出了重要贡献。



Robert Luke Devaney 波士顿大学教授，1973年获得加州大学伯克利分校博士学位，师从Stephen Smale。他是世界著名的动力系统和混沌理论专家。

本书译自原版*Differential Equations, Dynamical Systems, and an Introduction to Chaos*，并由Elsevier授权出版。



本书相关信息请访问：
图灵网站 http://www.turingbook.com
读者/作者热线：(010) 88593802
反馈/投稿/推荐信箱：contact@turingbook.com

分类建议 数学/微分方程

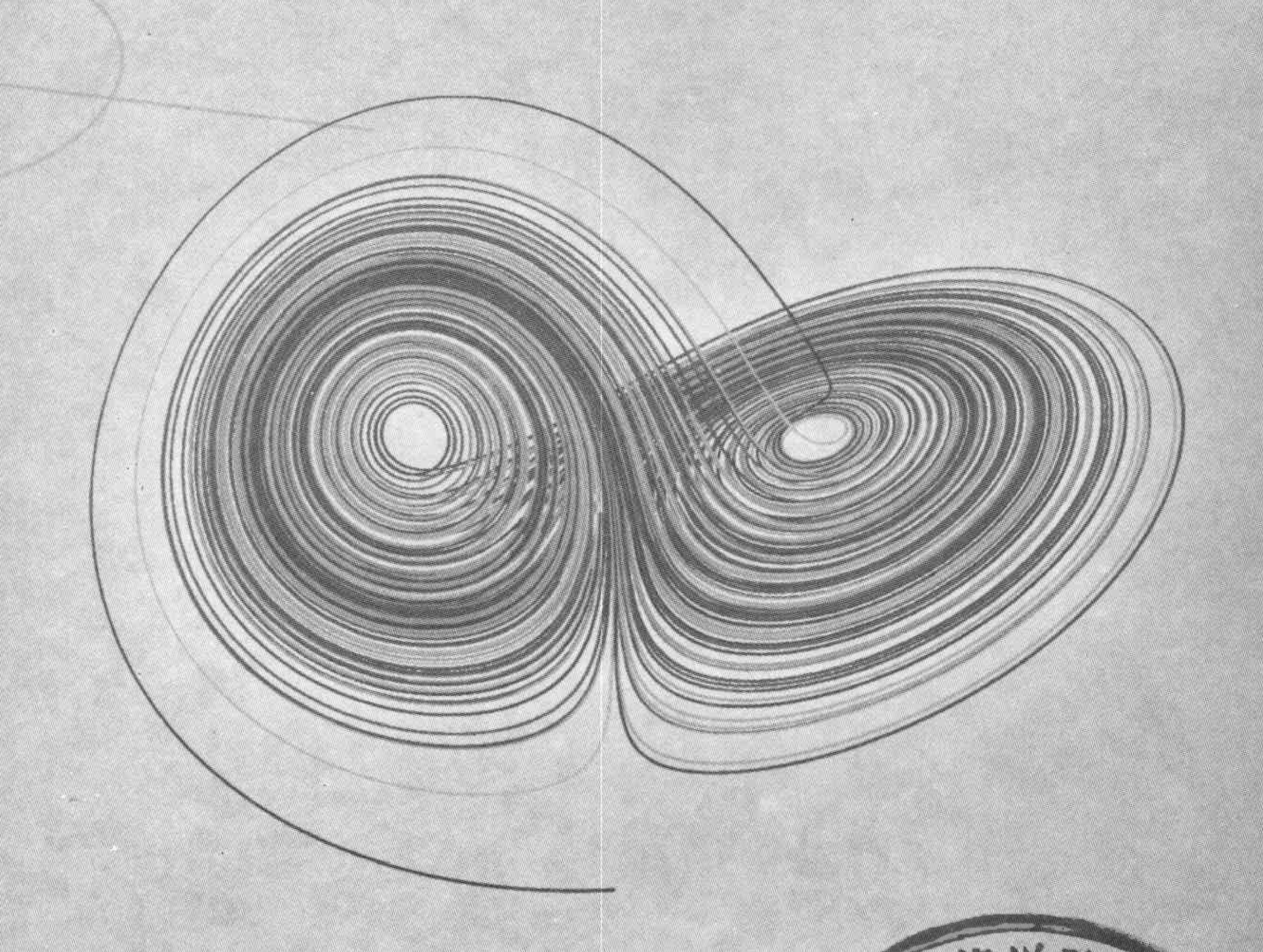
人民邮电出版社网址 www.ptpress.com.cn



ISBN 978-7-115-17218-1/O1
定价：59.00 元



图灵数学·统计学丛书 19



Differential Equations, Dynamical Systems, and an Introduction to Chaos

微分方程、动力系统与混沌导论

（第2版）

Morris W. Hirsch
［美］ Stephen Smale 著
Robert L. Devaney
甘少波 译

人民邮电出版社
北 京

图书在版编目（CIP）数据

微分方程、动力系统与混沌导论：第 2 版 /（美）赫希（Hirsch，M. W.），（美）斯梅尔（Smale，S.），（美）德瓦尼（Devaney，R. L.）著；甘少波译．—北京：人民邮电出版社，2008.4

（图灵数学·统计学丛书）

ISBN 978-7-115-17218-1

Ⅰ．微… Ⅱ．①赫…②斯…③德…④甘… Ⅲ．①微分方程—高等学校—教材②动力系统（数学）—高等学校—教材③混沌学—高等学校—教材 Ⅳ．O175 O415.5

中国版本图书馆 CIP 数据核字（2007）第 180218 号

内 容 提 要

本书主要介绍了线性系统、非线性系统、平面系统、洛伦茨系统、离散动力系统等，还分别介绍了这些知识在生物学、电路理论、力学等方面的应用．本书是介绍动力系统最值得推荐的一本经典教材，它从整体的、动态系统的观点介绍 ODE 的理论．每章后面都附有颇具参考价值的习题，可以帮助学生巩固所学知识．

本书适合作为高等院校数学专业以及其他理工科专业高年级学生学习动力系统的教材或参考书．

图灵数学·统计学丛书

微分方程、动力系统与混沌导论(第 2 版)

◆ 著　[美] Morris W. Hirsch　Stephen Smale　Robert L. Devaney
　译　甘少波
　责任编辑　明永玲

◆ 人民邮电出版社出版发行　北京市崇文区夕照寺街 14 号
　邮编　100061　电子函件　315@ptpress.com.cn
　网址　http://www.ptpress.com.cn
　北京铭成印刷有限公司印刷
　新华书店总店北京发行所经销

◆ 开本：700×1000　1/16
　印张：21.75
　字数：439 千字　2008 年 4 月第 1 版
　印数：1 – 4 000 册　2008 年 4 月北京第 1 次印刷

著作权合同登记号　图字：01-2007-1483 号

ISBN 978-7-115-17218-1/O1

定价：59.00 元

读者服务热线：(010)88593802　印装质量热线：(010)67129223

反盗版热线：(010)67171154

前　言

自本书第 1 版出版 30 年以来, 动力系统这一数学领域发生了巨大的变化. 在 20 世纪 70 年代早期, 我们很少有机会接触到高速计算机和计算机图形学. 在数学中也从来没有使用过“混沌”一词, 而且只有很少的数学家对微分方程和动力系统感兴趣.

这 30 年来情况发生了极大的变化. 计算机现在到处都是, 人们可以广泛地使用各种软件包来近似计算微分方程的解并将结果用图形显示出来. 如此一来, 分析非线性微分方程系统就变得比以往任何时候都更加容易. 诸如马蹄映射、同宿混乱和洛伦茨系统等复杂动力系统的发现以及数学上对它们的分析使得科学家们相信, 像平衡点或周期解这种简单的稳定运动不再总是微分方程解的最重要行为. 这些漂亮且触手可得的混沌现象促使许多学科的科学家和工程师们更加仔细地考察他们各自领域中的重要微分方程. 在许多情形下, 他们也在这些系统中发现了混沌行为. 如今, 几乎每个学科领域都出现了动力系统现象：从化学中的振荡 Belousov-Zhabotinsky 反应到电子工程中的混沌 Chua 电路, 从天体力学中的复杂运动到生态学中的分岔.

这样一来, 比起 20 世纪 70 年代, 微分方程和动力系统教材的读者就变得相当庞杂. 在这个版本中我们相应地做了几处结构调整, 这些变化包括：

(1) 线性代数的处理降低了难度. 我们删除了涉及抽象向量空间和赋范线性空间的一般理论, 并且不再完整地证明所有 $n \times n$ 矩阵都可以约化成标准形. 我们将主要处理大小不超过 4×4 的矩阵;

(2) 仔细地讨论了洛伦茨吸引子、Shil'nikov 系统以及双螺线吸引子中的混沌行为;

(3) 添加了许多新的应用, 同时对以前的应用也作了相应的更新;

(4) 专门用几章来讨论离散动力系统;

(5) 主要讨论 C^∞ 系统, 这样就简化了许多定理的假设.

本书由 3 个主要部分构成. 第一部分讨论线性微分方程系统和一些一阶非线性方程. 第二部分是该书的主要部分, 这里我们集中讨论非线性系统特别是二维系统, 以及这些系统在各种领域中的应用. 第三部分讨论高维系统, 这里, 我们着重研究平面系统所不具备的混沌行为, 并介绍研究这种行为的主要手段 —— 约化到一个离散动力系统.

为背景大不相同的读者来写一本书是一个很大的挑战. 我们认为本书可以作为微分方程后续课程的教材, 它的对象不仅仅是数学专业的学生, 而且还包括其他理工科学生, 本书将有助于他们培养出足够多的数学技巧来分析他们各自领域

中的各种微分方程. 使用本书的很多读者可能具有很强的线性代数和实分析背景, 而也有读者可能缺乏这些知识的训练. 为了使本书同时适合这两类读者, 我们开始先介绍了一些简单的低维微分方程系统. 对于有较强微分方程背景的读者来说, 这里大部分内容只不过是复习而已. 为了这些读者, 我们将一些新课题分散在本书的第一部分.

例如, 第 1 章讨论一阶方程. 在这一章的开始, 我们先讨论线性微分方程以及合理生物总量模型 (logistic population model), 只要了解微分方程的读者都应该熟悉这些课题. 复习完这些之后, 我们讨论带收割的合理模型 (logistic model), 这里既有常值收割也有周期收割. 这使得我们可以很快地引入分岔, 以及描述庞加莱映射和周期解. 这些课题通常不会出现在初等微分方程课程中, 但只要学过多元微积分就能很容易地理解这些内容. 当然, 背景有限的读者可以不妨跳过这些专门的课题, 而将注意力集中到更初等一些的内容.

第 2 章到第 6 章讨论线性微分方程系统. 同样地, 开始时很缓慢, 整个第 2 章和第 3 章都只是讨论平面微分方程系统和二维线性代数. 第 5 章和第 6 章介绍了高维线性系统, 然而我们的重点仍然只是三维和四维系统, 而不是最一般的 n 维系统, 虽然我们所描述的许多技巧都可以很容易地推广到更高维情形.

第二部分是本书的核心. 这里我们转而讨论非线性系统. 与线性系统不同, 讨论非线性系统会出现一些理论上的困难, 这些困难包括解的存在性和唯一性, 解对初始条件和参数的连续依赖性, 等等. 我们并不立刻就陷入这些困难的理论问题, 因为讨论这些问题要求相当坚实的实分析基础. 在第 7 章我们只是简单地陈述一些重要结果, 同时通过一些例子来解释这些定理说了些 (和没有说) 什么. 所有这些结果的证明都放在本书的最后一章.

在本书非线性部分开始的几章, 我们介绍了一些重要技巧, 这些技巧包括平衡点附近的线性化、零点集分析 (nullcline analysis)、稳定性、极限集以及分岔理论. 在这一部分的后面几章, 我们将这些想法运用到生物学、电子工程、力学等许多领域中产生的各种系统.

许多章的最后一节都是"探索". 这些节由一系列的问题和数值研究组成, 它们都是用来讨论与前面内容有关的一个特定课题或应用. 在每个探索中, 我们都简要地介绍了即将要做的课题, 并且提供了进一步了解这一课题的参考文献. 但是, 我们让读者利用前面所学的内容自己去讨论所得出的系统的行为. 我们常常为读者准备一系列的基础问题并经常提示如何去解决, 但是, 在许多情况下, 完全地分析这个系统可能会变成一个大的研究项目. 对于这些问题, 你不可能找到"书后面附的答案", 在许多情形下, 没有人知道完整答案. (当然, 你本人除外!)

本书的最后一部分专门讨论高维系统的复杂非线性行为, 即所谓的混沌行为. 我们通过著名的洛伦茨微分方程系统来引入这些思想. 在三维和更高维的情形, 我

们常常将理解微分方程复杂行为的问题约化成对一个离散动力系统或迭代映射的理解. 因而我们需要暂时转入离散系统的领域, 同时讨论怎样用符号动力学来完全描述某些混沌系统. 然后我们再回到非线性微分方程, 并将这些技巧运用到其他混沌系统, 其中包括出现同宿轨的系统.

我们有一个网站 http://www.math.bu.edu/hsd 专门用来发布与本书有关的消息. 在这里你可以找到本书的勘误表, 或给出你对本书的建议, 这里还有微分方程的教师和学生感兴趣的一些其他课题. 我们欢迎读者在这个站点上踊跃发表意见.

很高兴地感谢 Bard Ermentrout, John Guckenheimer, Tasso Kaper, Jerrold Marsden 和 Gareth Roberts 对本书第 1 版给出的许多好建议. 特别感谢 Daniel Look 和 Richard Moeckel, 他们仔细阅读了整个手稿. 本书中许多相平面图形的绘制都使用了优秀的 Mathematica 宏包 DynPac, 这是 Al Clark 为 Mathematica 写的一个动力系统宏包. 请浏览 http://www.me.rochester.edu/~clark/dynpac.html. 而且, 我的小狗 Killer Devaney 又啃光了所有手稿, 残留下来的所有错误都怨它!

致　谢

我们要感谢下面的审稿人：

明尼苏达大学的 Bruce Peckham
匹兹堡大学的 Bard Ermentrout
明尼苏达大学的 Richard Moeckel
加州理工学院的 Jerry Marsden
康奈尔大学的 John Guckenheimer
Holy Cross 学院的 Gareth Roberts
明尼苏达大学的 Rick Moeckel
加州大学圣迭戈分校的 Hans Lindblad
Gustavus Adolphus 学院的 Tom LoFaro
波士顿大学的 Daniel M. Look

目　录

第1章　一阶方程

本章的目的是提供一些初等但重要的一阶微分方程的例子，同时用这些例子在尽可能简单的情况下去解释常微分方程理论中的一些基本思想.

学过微分方程初级课程的读者应该熟悉本章的前几个例子，后几个例子，如带收割的合理模型，将使得读者对某些特定课题 (分岔、周期解和庞加莱映射) 有些感性认识. 在本书中，我们将经常讨论这些课题；在后面几章，我们将更加系统地讨论这些课题.

1.1　最简单的例子

学过微积分的学生都知道如下的微分方程

$$\frac{\mathrm{d}x}{\mathrm{d}t} = ax,$$

它是最简单的微分方程，同时也是最重要的微分方程之一. 首先，这个方程的含义是什么呢? 这里，$x = x(t)$ 是实变量 t 的实值未知函数，$\mathrm{d}x/\mathrm{d}t$ 是它的导数 (以后也用 x' 或 $x'(t)$ 表示导数). 另外，a 是一个参数，当 a 取不同的值时，就得到不同的微分方程. 方程告诉我们，对于 t 的任一取值，下面的关系成立：

$$x'(t) = ax(t).$$

该方程的解可由微积分得到. 设 k 为任一给定的实数，则函数 $x(t) = k\mathrm{e}^{at}$ 就是一个解，这是因为

$$x'(t) = ak\mathrm{e}^{at} = ax(t).$$

并且，*这个方程没有其他解*. 为看清这一点，设 $u(t)$ 为方程的任一解，计算 $u(t)\mathrm{e}^{-at}$ 的导数：

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(u(t)\mathrm{e}^{-at}\right) = u'(t)\mathrm{e}^{-at} + u(t)(-a\mathrm{e}^{-at}) = au(t)\mathrm{e}^{-at} - au(t)\mathrm{e}^{-at} = 0.$$

从而，$u(t)\mathrm{e}^{-at}$ 为一常数 k，于是 $u(t) = k\mathrm{e}^{at}$. 这就证明了我们的断言. 于是就找到了这个微分方程的所有可能解. 我们把一个微分方程的所有解的全体称为这个方程的**通解**.

当解在一个点 t_0 处的取值 u_0 确定后, 出现在解中的常数 k 就完全确定. 假设满足这个微分方程的函数 $x(t)$ 还要求满足 $x(t_0) = u_0$, 于是我们就必有 $ke^{at_0} = u_0$, 从而 $k = u_0e^{-at_0}$. 这样我们就确定了 k 的取值, 从而这个方程在满足特定**初始条件**$x(t_0) = u_0$ 时有唯一解. 为简单起见, 我们常取 $t_0 = 0$, 此时 $k = u_0$. 在上面取 $t_0 = 0$ 并不失一般性, 因为如果 $u(t)$ 是满足 $u(0) = u_0$ 的解, 则函数 $v(t) = u(t - t_0)$ 就是满足 $v(t_0) = u_0$ 的解.

通常将上面的问题写成**初值问题**的形式:

$$x' = ax, \qquad x(0) = u_0.$$

初值问题的一个解 $x(t)$ 不仅要满足微分方程, 而且还要在 $t = 0$ 处取特定的初值 u_0.

注意这个微分方程在 $k = 0$ 时有一个特别的解, 即常值解 $x(t) \equiv 0$. 像这样的常值解称为该方程的**平衡解**或**平衡点**. 平衡解通常都是微分方程最重要的解.

在方程 $x' = ax$ 中, 常数 a 可以看成一个参数, 当 a 变化时, 方程就变了, 当然解也要随之变化. 我们能否定性地描述解的变化呢? 这里 a 的符号将是关键:

(1) 如果 $a > 0$, 当 $k > 0$ 时, 极限 $\lim_{t\to\infty} ke^{at}$ 等于 ∞, 当 $k < 0$ 时, 该极限等于 $-\infty$;

(2) 如果 $a = 0$, ke^{at} 是常数;

(3) 如果 $a < 0$, $\lim_{t\to\infty} ke^{at} = 0$.

通过作解的简图可以形象地描述解的定性行为. 注意当 a 取正值和负值时, 解的行为相当不同. 当 $a > 0$ 时 (见图 1.1), 所有的非零解随着 t 增加都远离平衡点 (在 0 点), 而当 $a < 0$ 时 (见图 1.2), 所有的非零解随着 t 增加都趋于平衡点. 当平衡点附近的解都远离它时, 我们称该平衡点是一个**源点**; 而当附近的解都趋于它时, 该平衡点就称为一个**汇点**.

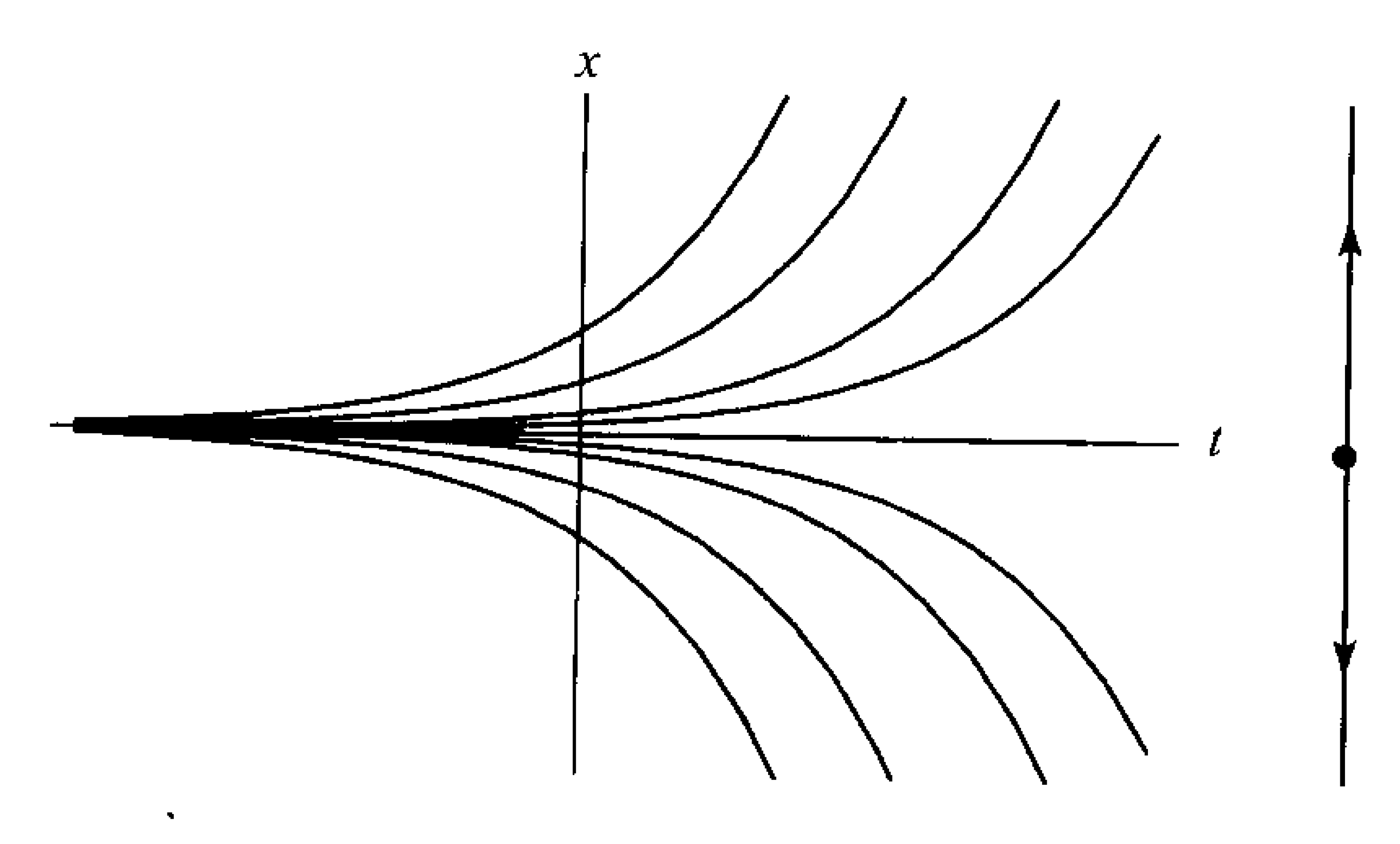


图 1.1 方程 $x' = ax$ 在 $a > 0$ 时解的图像和相线, 其中每一个函数图像都表示一个解

我们还可以将解画在**相线**(phase line) 上来描述解. 因为解 $x(t)$ 为时间的函数, 我们可以将 $x(t)$ 看成一个沿实直线运动的质点. 在平衡点处, 质点保持不动 (用一实心圆点表示), 而其他解则沿 x 轴上下运动, 在图 1.1 中用箭头表示.

在某种意义下, 当 $a \neq 0$ 时, 方程 $x' = ax$ 是**稳定的**. 精确地说, 当 a 用一个与之同号的 b 替换时, 解的定性行为不发生改变. 但当 $a = 0$ 时, a 的微小改变都将根本地改变解的行为. 于是, 我们说单参数微分方程族 $x' = ax$ 在 $a = 0$ 处出现了一个**分岔**.

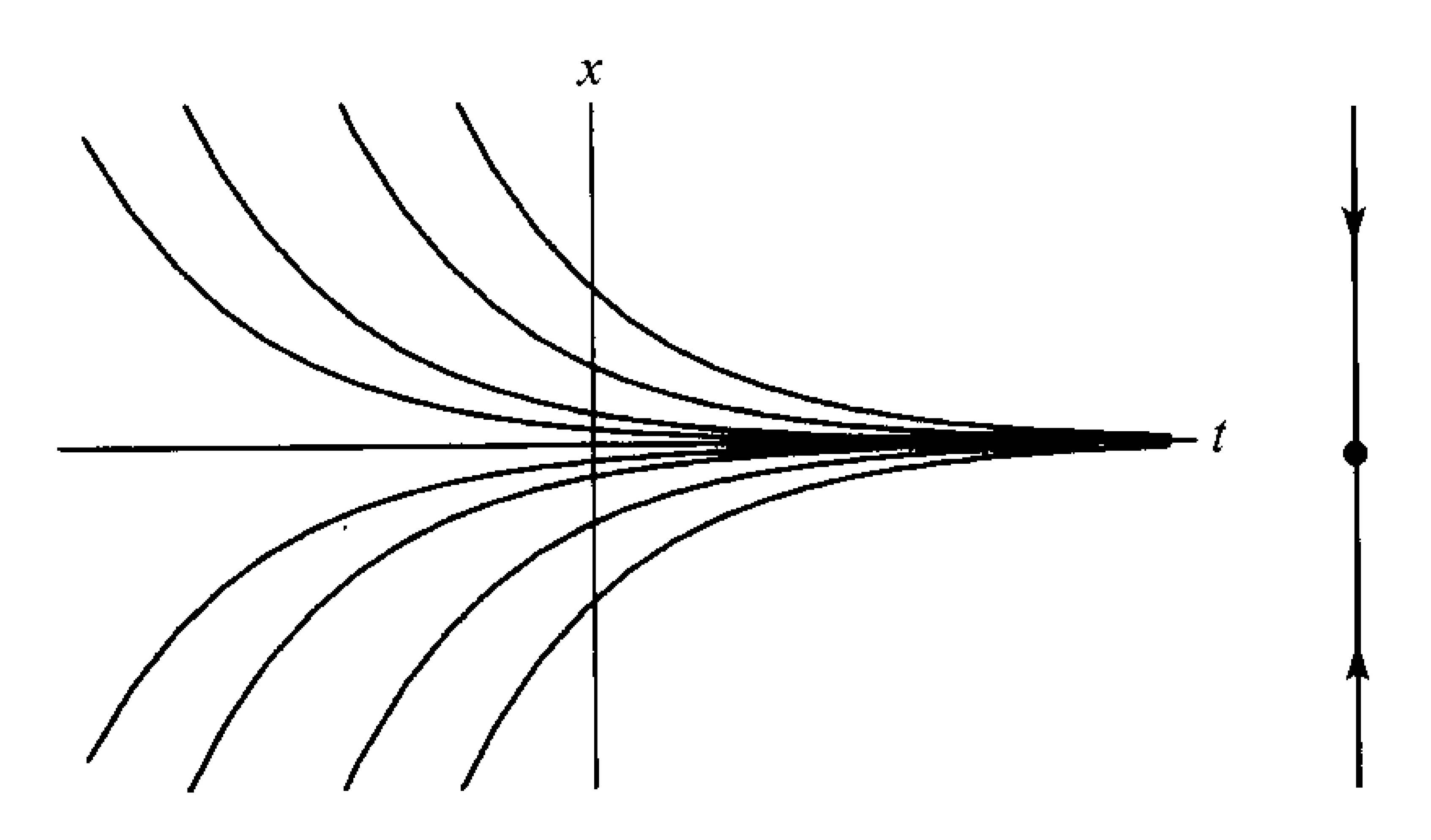


图 1.2 方程 $x' = ax$ 在 $a < 0$ 时解的图像和相线

1.2 合理的物种总量模型

前面的微分方程 $x' = ax(a > 0)$ 可以看成一个简单的物种总量增长模型. 变量 $x(t)$ 代表某个物种在时刻 t 的总量. 当我们假设总量的增长率 (即 $\mathrm{d}x/\mathrm{d}t$) 与总量成正比时, 就得到该微分方程. 当然, 这一简单的假设忽略了许多影响总量增长的因素, 例如, 实际的总量不可能无限制地增长.

为了在物种总量模型中将这种限制考虑进来, 我们进一步假设:

(1) 当总量较小时, 总量增长率几乎与总量成正比;

(2) 但当总量增长到很大时, 增长率就变成负的.

合理总量增长模型就是满足这些假设的一个微分方程, 其方程为

$$x' = ax\left(1 - \frac{x}{N}\right).$$

这里 a 和 N 为正参数: a 为 x 较小时的总量增长率, 而 N 则代表一种"理想"总量或"承载量". 注意, 当 x 较小时, 该微分方程本质上就是 $x' = ax$[因为 $1-(x/N) \approx 1$], 但当 $x > N$ 时, 则有 $x' < 0$. 从而这一简单方程满足上述的假设. 必须说明的是, 有许多其他的微分方程也满足上述假设, 而我们只是选取了差不多最简单的一个.

不失一般性, 我们假设 $N=1$, 也即选取单位使得承载量正好是 1 单位的总量, 而 $x(t)$ 则代表在 t 时刻的总量占理想总量的比例. 这样合理方程就简化成

$$x'=f_a(x)=ax(1-x).$$

这是一个一阶、自治、非线性微分方程的例子. 之所以称为**一阶**是因为在方程中只出现了 x 的一阶导数; 称为**自治**是因为方程的右端只与 x 有关, 而与时间 t 无关; 称为**非线性**是因为 $f_a(x)$ 是 x 的非线性函数. 前一个例子 $x'=ax$ 则是一阶、自治、线性微分方程.

合理微分方程的解可以很容易地用分离变量得到, 先分离变量再积分就有：

$$\int\frac{\mathrm{d}x}{x(1-x)}=\int a\ \mathrm{d}t.$$

然后将左端的被积函数写成部分分式：

$$\int\left(\frac{1}{x}+\frac{1}{1-x}\right)\mathrm{d}x.$$

分别对等式两端积分, 再求 x 就可以得到

$$x(t)=\frac{K\mathrm{e}^{at}}{1+K\mathrm{e}^{at}},$$

其中 K 是积分时产生的任意常数. 将 $t=0$ 代入上式可解得

$$K=\frac{x(0)}{1-x(0)}.$$

利用这个式子, 可以将解重写为

$$\frac{x(0)\mathrm{e}^{at}}{1-x(0)+x(0)\mathrm{e}^{at}}.$$

于是, 上述解的表达式对任意的初始总量 $x(0)$ 都成立. 当 $x(0)=1$ 时, 因为解为 $x(t)\equiv 1$, 我们得到一个平衡解. 类似地, $x(t)\equiv 0$ 也是一个平衡解.

于是, 我们就有了合理微分方程解的存在性. 但暂时我们还无法保证这就是方程的全部解, 当在第 7 章讨论微分方程解的存在唯一性问题时, 我们会继续研究这一问题.

为了对解的行为有一些定性认识, 我们画出该方程的**斜率场**. 方程的右端确定了解的图像在每一时刻 t 的斜率. 如图 1.3, 在 tx 平面上作一些很短的斜线, 这些短斜线在 (t,x) 点处的斜率为 $ax(1-x)$. 解的图像必须与这个斜率场处处相切. 从图上易见, 所有对应于初值 $x(0)>0$ 的解都趋向于理想总量 $x(t)\equiv 1$, 这正好与

我们的假设吻合. 当初值 $x(0) < 0$ 时, 解将趋于 $-\infty$. 当然这些解在总量模型中是没有意义的.

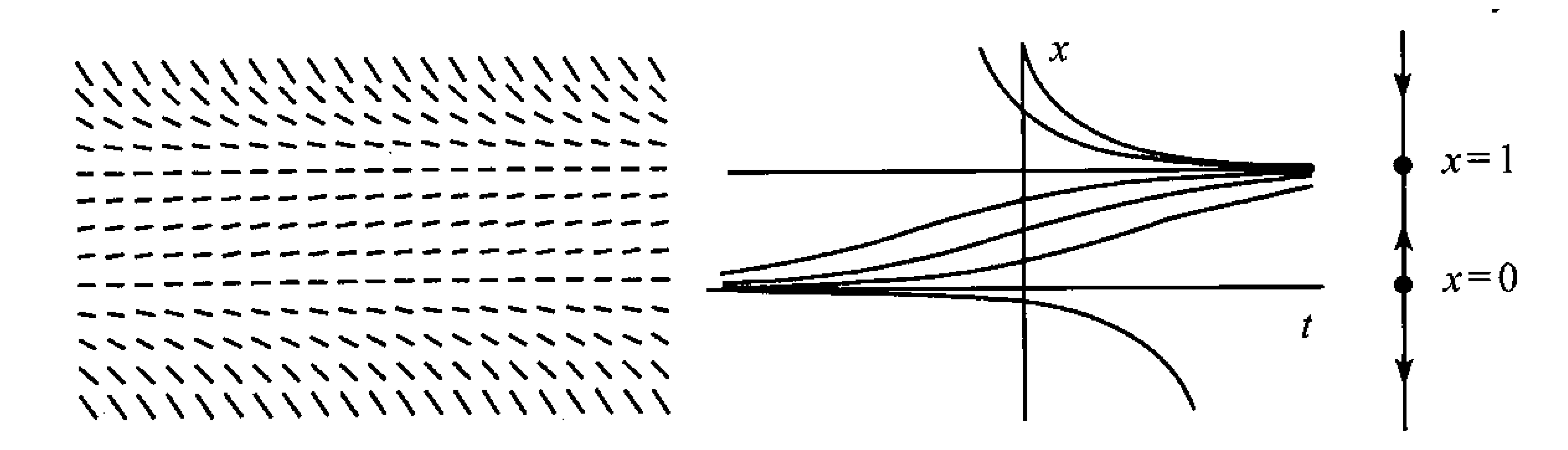


图 1.3 方程 $x' = ax(1-x)$ 的斜率场、解的图像和相线

注意, 我们还可以从函数 $f_a(x) = ax(1-x)$ 的图像上认识这些行为. 该函数图像 (见图 1.4) 与 x 轴交于 $x=0$ 和 $x=1$ 两个点, 这正对应于两个平衡点. 当 $0 < x < 1$ 时, $f_a(x) > 0$, 从而在任何满足 $0 < x < 1$ 的 (t, x) 点处, 斜率为正数, 从而解在这个区域将增加. 而在 $x < 0$ 或 $x > 1$ 时, $f_a(x) < 0$, 故解将减小, 正如图 1.3 的解的图像和相线所示.

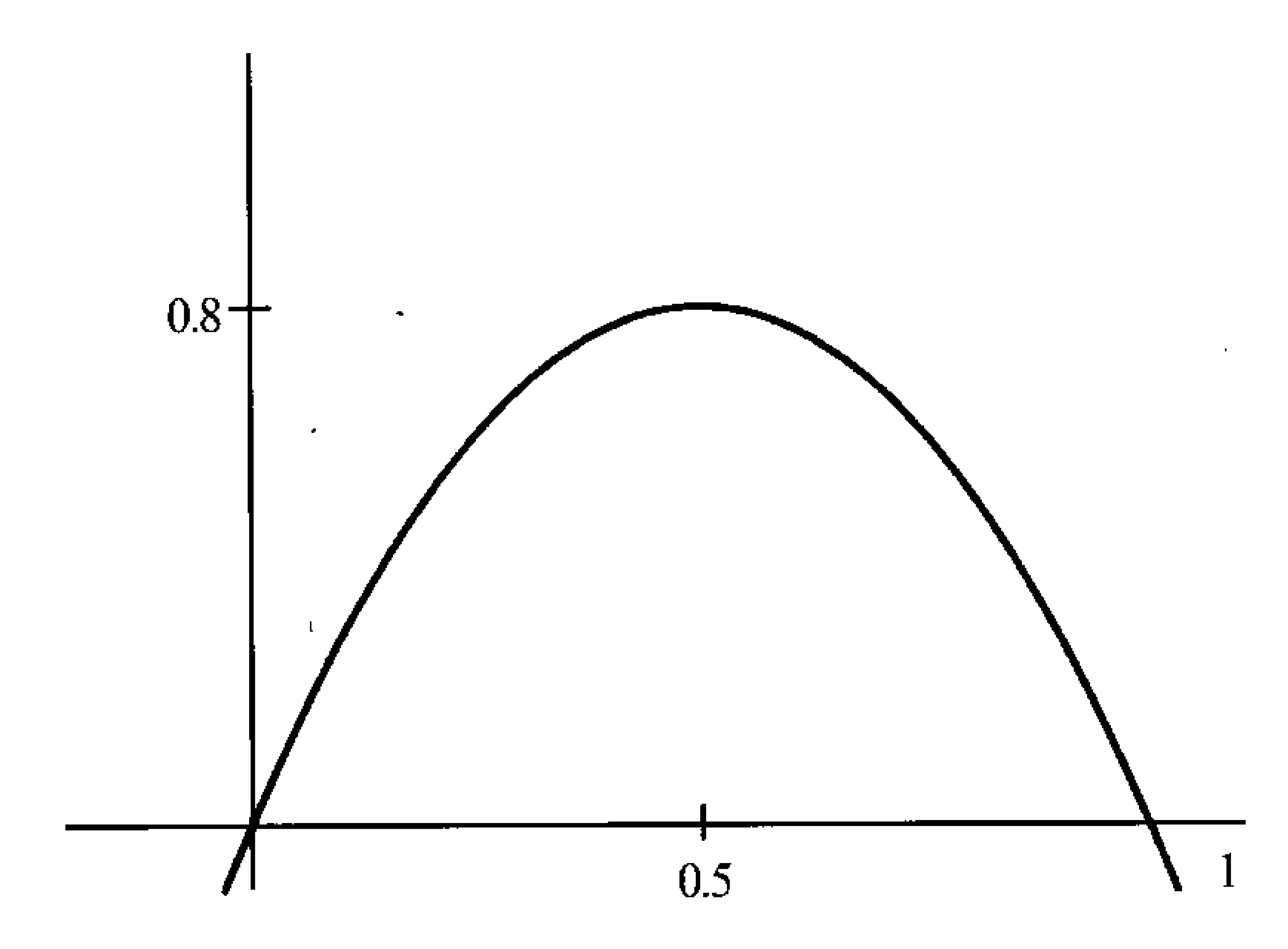


图 1.4 当 $a = 3.2$ 时, 函数 $f_a(x) = ax(1-x)$ 的图像

同样, 我们还可以从 f_a 的图像上看出：$x=0$ 是一源点而 $x=1$ 是一汇点. 在 $x=0$ 附近, 当 $x>0$ 时, $f_a(x) > 0$, 斜率为正, 解增加; 但当 $x<0$ 时, $f_a(x) < 0$, 斜率为负, 解减小. 这意味着附近的解都要远离 0, 即 0 为源点. 同样, 1 为汇点.

此外, 我们还可以通过分析得到这些信息. 由于 $f_a'(x) = a - 2ax$, 从而有 $f_a'(0) = a > 0$ 而 $f_a'(1) = -a < 0$. 由于 $f_a'(0) > 0$, 当 x 通过 0 时, 斜率将单调增加, 于是在 $x=0$ 的下方, 斜率取负值, 而在 $x=0$ 的上方, 斜率取正值. 因而, 解要远离 $x=0$. 同样, $f_a'(1) < 0$ 将使得解趋于 $x=1$, 从而使得这个平衡点成为一汇点. 在今后, 我们会遇到很多这样通过计算导数来确定平衡点附近解的定性行为的例子.

例　为了进一步解释这种定性分析的思想, 我们来考虑微分方程

$$x' = g(x) = x - x^3.$$

该方程有 3 个平衡点, $x = 0, \pm 1$. 由于 $g'(x) = 1 - 3x^2$, 所以 $g'(0) = 1$, 从而平衡点 0 为一源点. 同样, $g'(\pm 1) = -2$, 所以在 ± 1 处的平衡点为汇点. 在这些平衡点之间的斜率场非零. 利用这些信息, 我们可以立刻作出方程的相线 (见图 1.5).　■

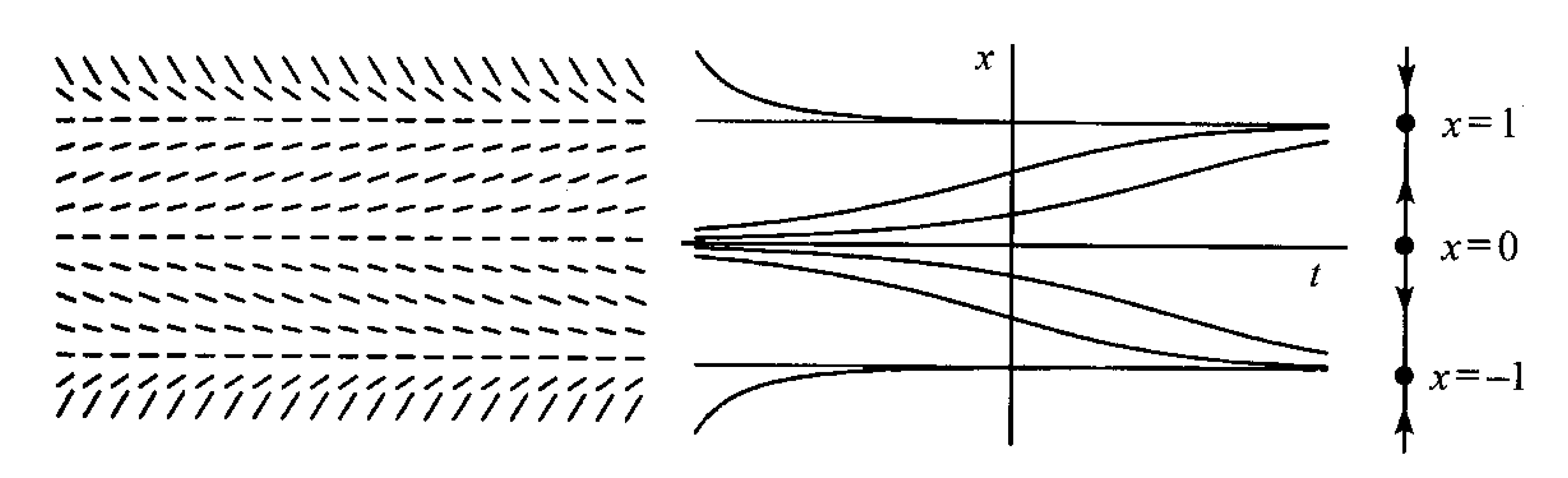


图 1.5　方程 $x' = x - x^3$ 的斜率场、解的图像和相线

1.3　常值收割与分岔

现在将物种的收割考虑进来以修改合理模型. 假设物种遵循参数 $a = 1$ 时的合理假设, 但它们同时以常速率 h 被收割. 此时, 微分方程为

$$x' = x(1 - x) - h,$$

其中 $h \geqslant 0$ 为一新参数.

下面我们不解方程 (该方程可以解出, 见本章的习题 6) 而直接利用函数

$$f_h(x) = x(1 - x) - h$$

的图像来 “读出” 解的定性行为. 在 $0 < h < 1/4, h = 1/4$ 及 $h > 1/4$ 三种不同的情况下, 图 1.6 作出 f_h 的图像. 直接计算可得：当 $0 \leqslant h < 1/4$ 时, f_h 有两个根; 当 $h = 1/4$ 时有一个根; 而当 $h > 1/4$ 时, 没有根 (如图所示). 从而, 在 $0 < h < 1/4$ 时, 该微分方程有两个平衡点 x_l 和 x_r, $0 \leqslant x_l < x_r$. 容易验证 $f_h'(x_l) > 0$, 于是 x_l 是一源点, 而 $f_h'(x_r) < 0$, 于是 x_r 是一汇点.

当 h 通过 $h = 1/4$ 时, 另一种分岔现象发生了：当 h 单调增加通过 $1/4$ 时, 两个平衡点 x_l 和 x_r 重合, 而当 $h > 1/4$ 时, 平衡点消失. 事实上, 当 $h > 1/4$ 时, 对所有的 x, 总有 $f_h(x) < 0$. 数学上, 这意味着在将来微分方程的解都要递减到 $-\infty$.

我们用**分岔图**来形象地记录这些变化. 在该图中, 我们用 h 代表横坐标, 而对应于每一个 h 值, 画出相应的相线. 图中的曲线代表每个 h 值所对应的平衡点. 这

让我们从另一角度看到, 当 h 通过 1/4 时, 源点和汇点融合成一个平衡点, 然后消失 (见图 1.7).

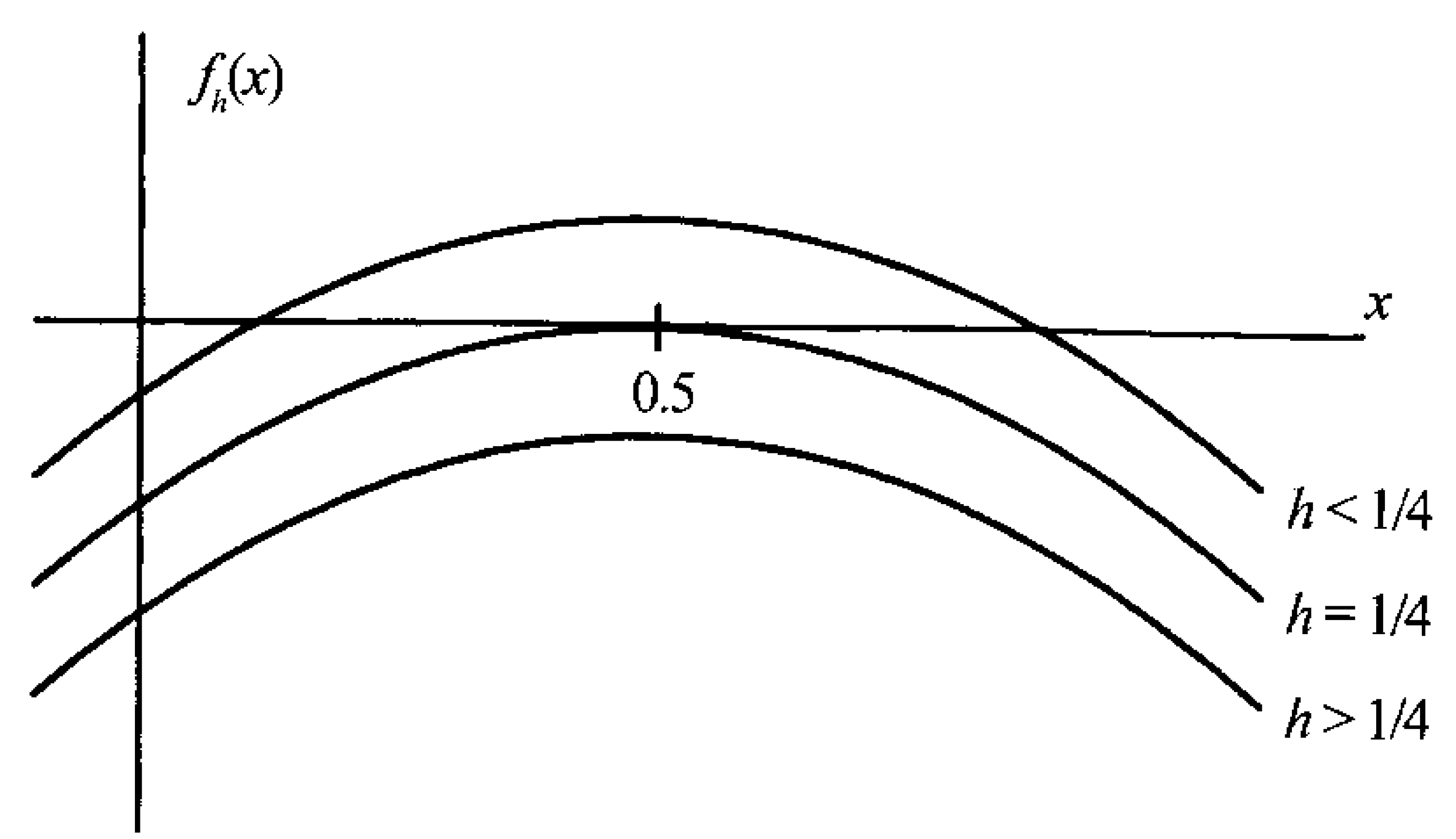


图 1.6 函数 $f_h(x) = x(1-x) - h$ 的图像

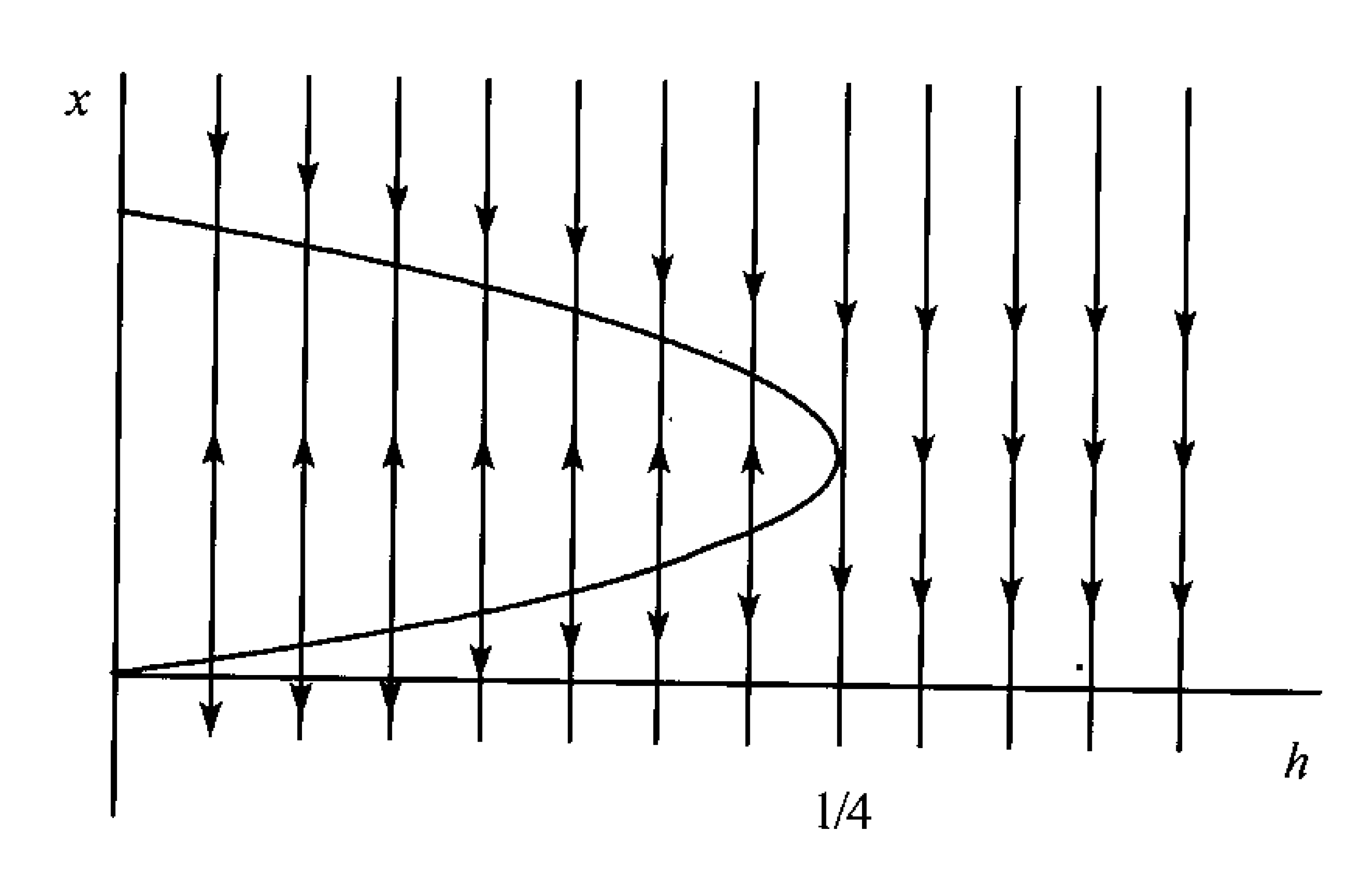


图 1.7 $f_h(x) = x(1-x) - h$ 的分岔图

在生态学上, 这种分岔对应于所研究物种的灾难. 当收割率为 1/4 或更低时, 只要初始总量充分大 ($x(0) \geqslant x_l$), 总量就能保持. 当 $h = 1/4$ 时, 收割率的微小变化都将要导致物种命运的大变化, 例如, 只要收割率 $h > 1/4$, 该物种就要灭绝.

上述现象突出地说明了在微分方程族中觉察到分岔是多么地重要, 在后续章节中, 我们还会经常遇到这种情况. 必须说明的是, 上面的物种总量模型虽然简单, 但它预测收割率微小变化所导致的物种灾难性变化已经在现实中被多次观察到.

例 我们来看另外一种分岔现象. 为此, 我们考虑微分方程族

$$x' = g_a(x) = x^2 - ax = x(x-a),$$

这里 a 为参数. 该方程的平衡点为 $x = 0$ 和 $x = a$. 计算可得 $g_a'(0) = -a$, 从而当 $a > 0$ 时, 0 是一汇点, 而当 $a < 0$ 时, 0 是一源点. 同样, $g_a'(a) = a$, 从而当 $a < 0$ 时, $x = a$ 为汇点, 而当 $a > 0$ 时为源点. 在 $a = 0$ 时出现了分岔, 因为此时, 方程只有一个平衡点. 进一步, 当 a 增加地通过 0 时, 在 $x = 0$ 处的平衡点从源点变为汇点; 与此同时, 在 $x = a$ 处的平衡点从汇点变为源点. 这一分岔图如图 1.8 所示. ■

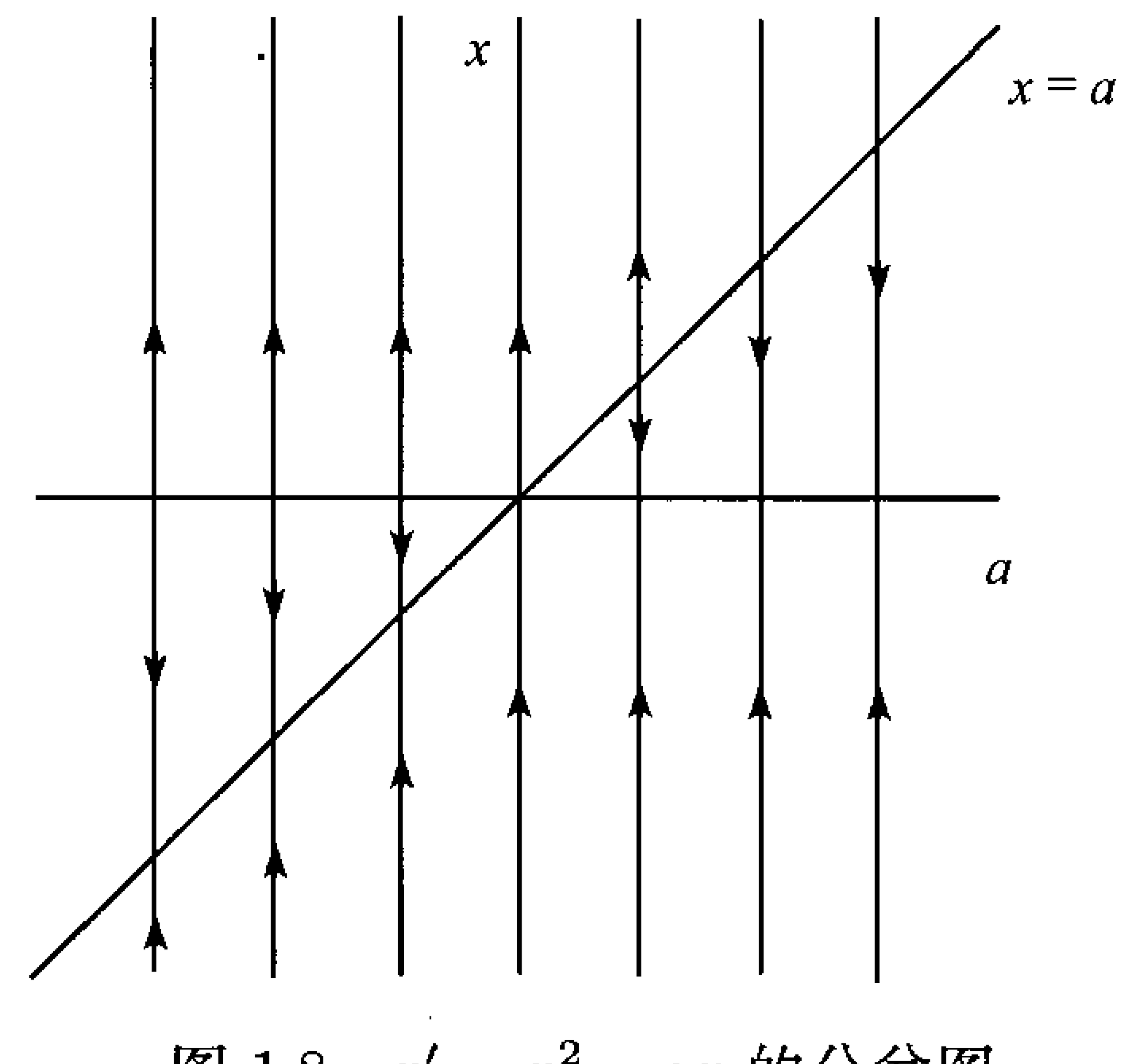


图 1.8 $x' = x^2 - ax$ 的分岔图

1.4 周期收割与周期解

现在不再假设在合理模型中的收割率为常数. 例如, 许多种类的鱼群在温暖季节的收割率比在寒冷月份的收割率要高. 于是, 假设物种的收割率是周期变化的.

$$x' = f(t,x) = ax(1-x) - h(1+\sin(2\pi t))$$

就是这样的一个模型, 其中 a 和 h 都是正参数 (见图 1.9). 收割率在时刻 $t = \frac{1}{4} + n$ (n 代表年份, n 为整数) 时达到最大值 $-2h$; 而正好在半年后, 即 $t = \frac{3}{4} + n$ 时达到最小值 0. 注意该微分方程显式地依赖于时间, 从而这是一个**非自治**的微分方程的例子. 如同在自治情形一样, 该方程的一个解 $x(t)$ 必须对所有的 t 都满足 $x'(t) = f(t, x(t))$. 此外, 该方程不再是可分离变量的, 我们无法用通常微积分的方法求出解的解析表达式, 这迫使我们去采用更加定性化的方法.

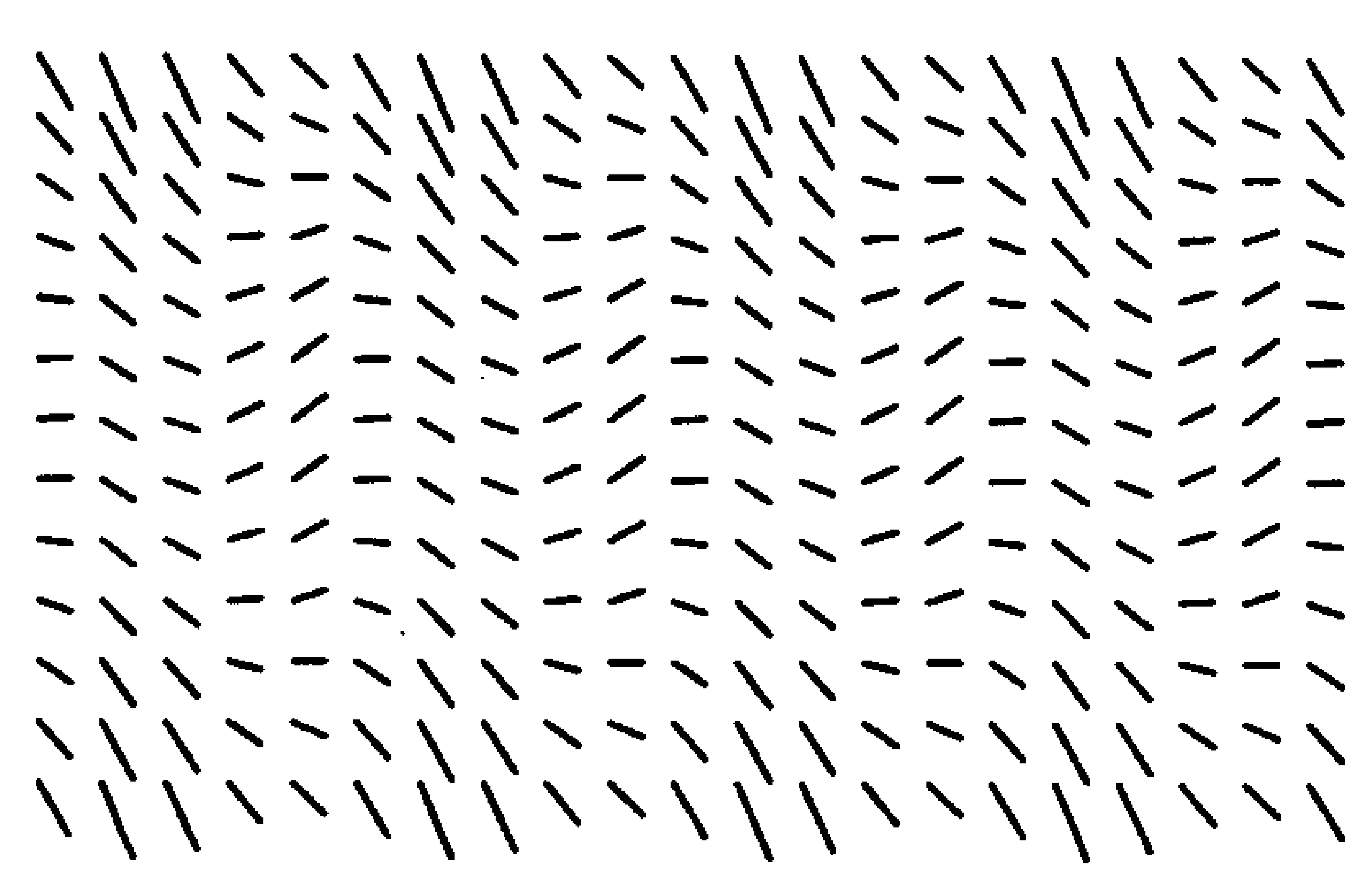


图 1.9 $x' = x(1-x) - h(1+\sin(2\pi t))$ 的斜率场

为了描述这种情况下总量的变化趋势, 首先注意到, 该微分方程的右端对时间变量是以 1 为周期的, 即 $f(t+1,x)=f(t,x)$. 这一事实多少可以简化求解问题. 假设我们知道了所有初值问题在 $0\leqslant t\leqslant 1$ 时 (不是所有时间!) 的解, 则事实上我们就知道了任意时刻的解. 假设 $x_1(t)$ 是定义在 $0\leqslant t\leqslant 1$ 上的一个解, 满足 $x_1(0)=x_0$, 而 $x_2(t)$ 为满足 $x_2(0)=x_1(1)$ 的一个解, 则可以通过定义 $x_1(t+1)=x_2(t)(0\leqslant t\leqslant 1)$ 将解 $x_1(t)$ 进行延伸. 由于

$$x_1'(t+1)=x_2'(t)=f(t,x_2(t))=f(t+1,x_1(t+1)),$$

因而延伸后的函数仍然是一个解. 从而, 一旦知道了解在区间 $0\leqslant t\leqslant 1$ 上的行为, 我们就可以用同样的方法得到解在所有时间区间的行为.

其次, 假设知道了满足任一初始条件 $x(0)=x_0$ 的解在时刻 $t=1$ 时的取值 $x(1)$, 则对于每一个初值 x_0, 都对应于满足 $x(0)=x_0$ 的解 $x(t)$ 的一个值 $x(1)$. 这样就定义了一个函数 $p(x_0)=x(1)$. 将这个函数与自身作一次复合, 就得到了以 x_0 为初值的解在 $t=2$ 处的取值, 即 $p(p(x_0))=x(2)$. 将这个函数与自身作 n 次复合, 就可以算出该解在时刻 n 的值, 从而也就知道了该解的变化趋势.

上述定义的函数 p 称为该微分方程的**庞加莱映射**. 有了这样一个函数, 我们就可以从连续动力系统 (微分方程) 的领域转化到较易理解的离散动力系统的领域 (迭代函数). 例如, 假设对某一个初值 x_0 有 $p(x_0)=x_0$, 即 x_0 为函数 p 的**不动点**. 则从前面的观察可知, 对每一整数 n, 都有 $x(n)=x_0$, 进一步, 对于满足 $0<t<1$ 的任何 t, 也有 $x(t)=x(t+1)$, 从而对于所有的整数 n 都有 $x(t+n)=x(t)$. 这说明满足初始条件 $x(0)=x_0$ 的解关于 t 是一个以 1 为周期的周期函数. 这样的解称为微分方程的**周期解**. 在图 1.10 中, 我们画出了具有周期收割的合理模型的几个周期解. 注意到满足初值条件 $x(0)=x_0$ 的解是一个周期解, 从而有 $x_0=p(x_0)=p((x_0))\cdots$. 同样, 满足初值条件 $x(0)=\hat{x}_0$ 的解也是一个周期解, 从而也有 $p(\hat{x}_0)=\hat{x}_0$.

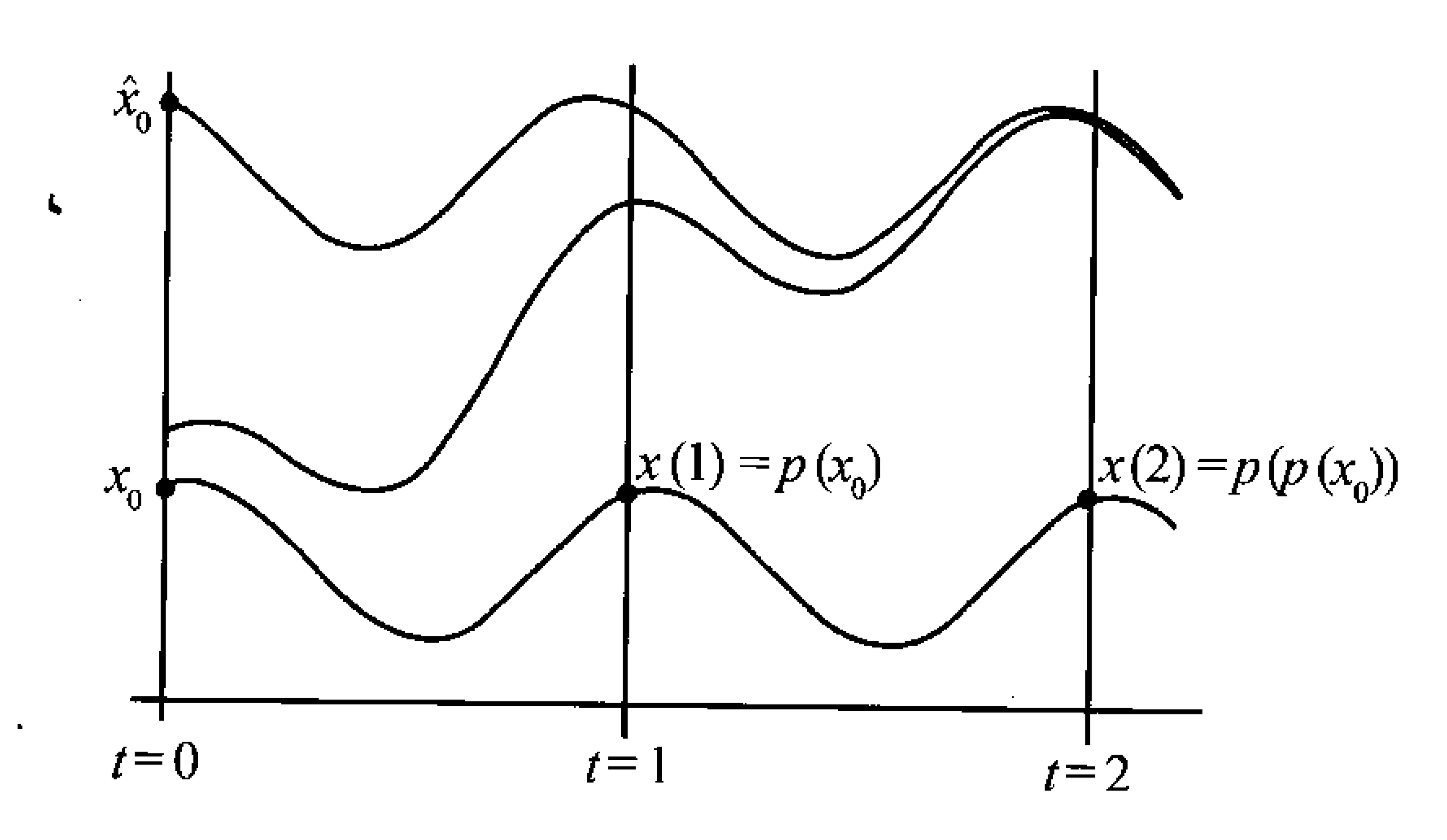


图 1.10 $x'=5x(1-x)-0.8(1+\sin(2\pi t))$ 的庞加莱映射

不幸的是, 计算微分方程的庞加莱映射通常是一件不可能的事. 所幸对于具有周期收割的合理方程, 我们可以做到.

1.5 计算庞加莱映射

在计算该方程的庞加莱映射之前, 我们先引入一些重要的术语. 为了强调解对初值 x_0 的依赖关系, 将相应的解记为 $\phi(t,x_0)$. 函数 $\phi:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ 称为该微分方程诱导的**流**. 如果将变量 x_0 固定, 则函数

$$t\to\phi(t,x_0)$$

正是该微分方程对应于初值 x_0 的解的另一种表述方式. 有时我们也将这个函数记为 $\phi_t(x_0)$.

例 对于前面提到的第一个例子, $x'=ax$, 它的流为

$$\phi(t,x_0)=x_0\mathrm{e}^{at}.$$

对于合理方程 (无收割的情形), 流为

$$\phi(t,x_0)=\frac{x(0)\mathrm{e}^{at}}{1-x(0)+x(0)\mathrm{e}^{at}}.$$

现在我们回到带周期收割的合理方程

$$x'=f(t,x)=ax(1-x)-h(1+\sin(2\pi t)).$$

该方程满足初值条件 $x(0)=x_0$ 的解由 $t\to\phi(t,x_0)$ 给出. 虽然我们不知道具体的表达式, 但根据微积分基本定理, 以及

$$\frac{\partial\phi}{\partial t}(t,x_0)=f(t,\phi(t,x_0)),\qquad \phi(0,x_0)=x_0,$$

可知这个解满足

$$\phi(t,x_0)=x_0+\int_0^t f(s,\phi(s,x_0))\mathrm{d}s.$$

将上式对 x_0 求导, 由链式法则可得

$$\frac{\partial\phi}{\partial x_0}(t,x_0)=1+\int_0^t\frac{\partial f}{\partial\phi}(s,\phi(s,x_0))\cdot\frac{\partial\phi}{\partial x_0}(s,x_0)\mathrm{d}s.$$

现在, 令

$$z(t)=\frac{\partial\phi}{\partial x_0}(t,x_0).$$

注意到

$$z(0)=\frac{\partial\phi}{\partial x_0}(0,x_0)=1,$$

将 z 对 t 求导可得

$$z'(t)=\frac{\partial f}{\partial\phi}(t,\phi(t,x_0))\cdot\frac{\partial\phi}{\partial x_0}(t,x_0)=\frac{\partial f}{\partial\phi}(t,\phi(t,x_0))\cdot z(t).$$

虽然我们还是不能显式地知道 $\phi(t,x_0)$, 但上一方程却告诉我们 $z(t)$ 满足微分方程

$$z'(t)=\frac{\partial f}{\partial\phi}(t,\phi(t,x_0))z(t)$$

以及 $z(0)=1$. 于是, 通过分离变量, 可求得该方程的解为

$$z(t)=\exp\int_0^t\frac{\partial f}{\partial\phi}(s,\phi(s,x_0))\mathrm{d}s,$$

从而可得

$$\frac{\partial\phi}{\partial x_0}(1,x_0)=\exp\int_0^1\frac{\partial f}{\partial\phi}(s,\phi(s,x_0))\mathrm{d}s.$$

由于 $p(x_0)=\phi(1,x_0)$, 这样我们就确定了庞加莱映射的导数 $p'(x_0)$. 注意到 $p'(x_0)>0$, 故 p 为单调增加函数.

再微分一次可得

$$p''(x_0)=p'(x_0)\left(\int_0^1\frac{\partial^2 f}{\partial\phi\partial\phi}(s,\phi(s,x_0))\cdot\exp\left(\int_0^s\frac{\partial f}{\partial\phi}(u,\phi(u,x_0))\mathrm{d}u\right)\mathrm{d}s\right),$$

这个式子看起来很可怕, 但是由于

$$f(t,x_0)=ax_0(1-x_0)-h(1+\sin(2\pi t)),$$

因而

$$\frac{\partial^2 f}{\partial\phi\partial\phi}\equiv-2a,$$

故可得 $p''(x_0)<0$. 于是, 庞加莱映射的图像是向下凹的. 这蕴涵 p 的图像至多穿过对角线 $y=x$ 两次, 即 x 至多只有两个值满足 $p(x)=x$. 从而庞加莱映射至多只有两个不动点. 这些不动点对应于原微分方程的周期解, 而且这些解满足, 对所有的 t, $x(t+1)=x(t)$. 我们还可以这样说, 当初值 x_0 为这些不动点之一时, 流

$\phi(t,x_0)$ 关于 t 是以 1 为周期的周期函数. 在 $h=0.8$ 的特例时, 在图 1.10 中, 我们看见了两个这样的解. 在图 1.11 中, 又有两个解看起来是周期的. 注意其中一个周期解看起来吸引附近所有的解, 而另一个排斥附近所有的解. 在本书的后面, 我们还会经常讨论这些概念, 并使得它们更精确.

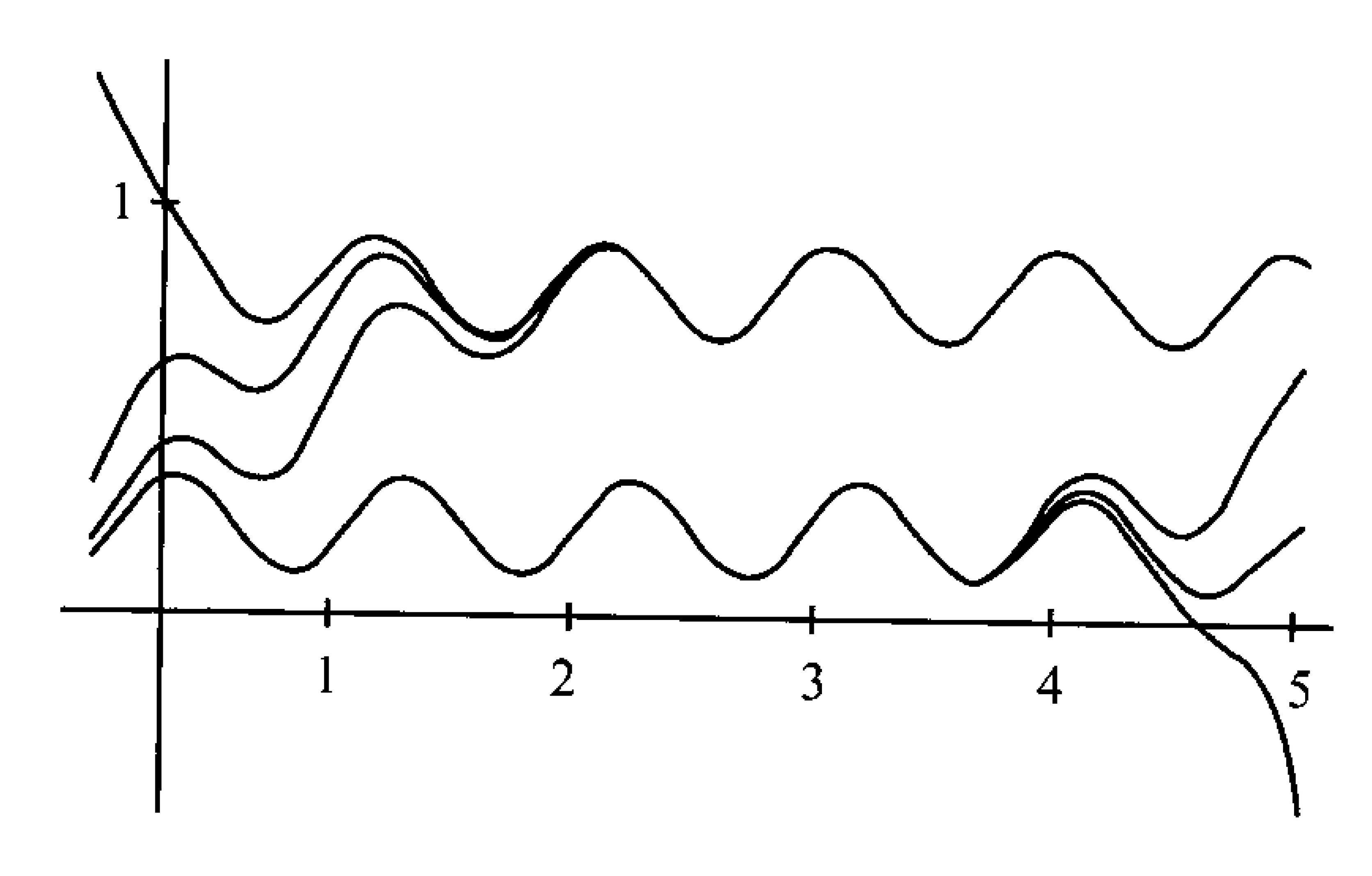


图 1.11　$x'=5x(1-x)-0.8(1+\sin(2\pi t))$ 的几个解

回想一下, 该微分方程还依赖于收割参数 h. 如图 1.11 所示, 当 h 比较小时, 有两个不动点. 将 f 对 h 微分可得

$$\frac{\partial f}{\partial h}(t,x_0)=-(1+\sin 2\pi t),$$

于是, $\partial f/\partial h<0$ (除去 $t=3/4$ 的情形). 这蕴涵在任一点 (t,x_0) 处的斜率随着 h 的增加而减小. 从而, 庞加莱映射的值也随着 h 的增加而减小. 从而 h 有唯一的取值 h_*, 使得在该点处庞加莱映射恰好只有一个不动点. 而当 $h>h_*$ 时, p 没有不动点, 此时对所有的初值都有 $p(x_0)<x_0$. 这意味着该物种也将灭绝. ■

1.6　探索: 一个双参数族

考虑依赖于两个参数 a 和 b 的微分方程族

$$x'=f_{a,b}(x)=ax-x^3-b.$$

这个探索的目的是, 综合本章提到的各种方法, 在二维参数平面 (ab 平面) 上作出该微分方程的完整图像. 可以先对该微分方程在计算机上作实验, 但需要严格验算你的观察结果.

(1) 首先固定 $a=1$, 利用 $f_{1,b}$ 的图像, 对依赖于 b 的微分方程族建立分岔图.

(2) 对 $a=0$ 以及 $a=-1$ 重复上一步.

(3) 其他的 a 值对应的分岔图是什么?

(4) 然后固定 b, 利用函数图像, 这次, 对依赖于 a 的微分方程族建立分岔图.

(5) 在 ab 平面上, 根据微分方程平衡点数目的不同, 勾画出相应的区域, 注意要画出这些区域的边界.

(6) 利用相线和 $f_{a,b}(x)$ 的图像, 描述当参数通过边界时出现的分岔.

(7) 详细描述在 $a=b=0$ 时, 当 a 和 (或)b 变化时出现的分岔.

(8) 考虑当 $|b|$ 很小时的微分方程 $x'=x-x^3-b\sin(2\pi t)$. 该如何描述这个方程的解? 它有周期解吗?

(9) 实验上, 当 $|b|$ 增加时会出现什么? 观察到分岔了吗? 试解释你观察到的现象.

习　题

1. 找出微分方程 $x'=ax+3$ 的通解, 其中 a 为参数. 该方程有哪些平衡点? 对 a 的哪些取值, 平衡点是汇点? 又对 a 的哪些取值, 平衡点是源点?
2. 对下列每个微分方程, 找出所有的平衡点, 并判定它们是汇点、源点, 还是都不是, 同时作出相线的简图.
   (a) $x'=x^3-3x$;　(b) $x'=x^4-x^2$;　(c) $x'=\cos x$;
   (d) $x'=\sin^2 x$;　(e) $x'=|1-x^2|$.
3. 下列微分方程族都依赖于参数 a, 画出其分岔图的简图.
   (a) $x'=x^2-ax$;　(b) $x'=x^3-ax$;　(c) $x'=x^3-x+a$.
4. 考虑如图 1.12 所示的函数 $f(x)$.
   (a) 画出微分方程 $x'=f(x)$ 相线的简图.
   (b) 令 $g_a(x)=f(x)+a$. 画出对应微分方程族 $x'=g_a(x)$ 分岔图的简图.
   (c) 描述这个微分方程族中出现的各种分岔.

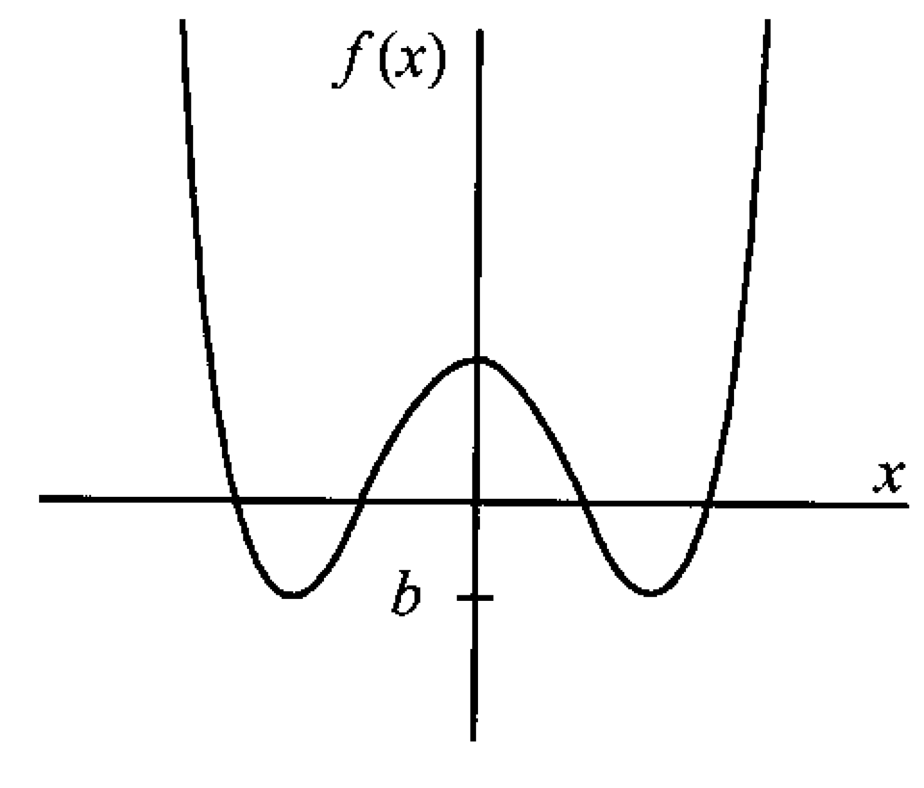


图1.12　函数f的图像

5. 考虑微分方程族

$$x'=ax+\sin x,$$

其中 a 为参数.
   (a) 画出 $a=0$ 时相线的简图.
   (b) 利用 ax 和 $\sin x$ 的图像, 确定在 a 从 -1 增加到 1 时出现的所有分岔的定性行为.
   (c) 画出该微分方程族分岔图的简图.
6. 对所有的参数值 $h>0$, 找出带常值收割的合理微分方程

$$x'=x(1-x)-h$$

的通解.

7. 考虑非自治微分方程

$$x' = \begin{cases} x-4, & \text{如果 } t < 5, \\ 2-x, & \text{如果 } t \geqslant 5. \end{cases}$$

(a) 找出该方程满足 $x(0)=4$ 的一个解. 描述这个解的定性行为.

(b) 找出该方程满足 $x(0)=3$ 的一个解. 描述这个解的定性行为.

(c) 描述这个方程的任一解在 $t \to \infty$ 时的定性行为.

8. 考虑形如 $x' = ax + f(t)$ 的一阶线性微分方程, 其中 $a \in \mathbb{R}$. 令 $y(t)$ 为该方程的任一解. 证明该方程的通解为 $y(t) + c\exp(at)$, 其中 $c \in \mathbb{R}$ 为任意常数.

9. 考虑形如 $x' = a(t)x$ 的一阶线性非自治微分方程.

(a) 找出该系统解的积分表达式.

(b) 证明你所找到的表达式为该系统的通解.

10. 考虑微分方程 $x' = x + \cos t$.

(a) 找出该方程的通解.

(b) 证明该方程有唯一的周期解.

(c) 计算该方程的庞加莱映射 $p: \{t=0\} \to \{t=2\pi\}$ 并利用它重新验证该方程有唯一周期解.

11. 一阶微分方程的解不必对所有时间都有定义.

(a) 找出方程 $x' = x^2$ 的通解.

(b) 讨论解的定义域.

(c) 给出一个微分方程的例子, 使得它满足 $x(0)=0$ 的解只在 $-1 < t < 1$ 上有定义.

12. 一阶微分方程满足给定初值条件的解并不总是唯一的.

(a) 证明微分方程 $x' = x^{1/3}$ 满足 $x(0)=0$ 的解有无穷多个.

(b) 讨论方程 $x' = x/t, x(0) = x_0$ 的解的相应情况.

(c) 讨论方程 $x' = x/t^2, x(0) = 0$ 的解的情况.

13. 设 x_0 为一阶自治微分方程 $x' = f(x)$ 的平衡点.

(a) 假设 $f'(x_0) = 0$. x_0 附近的解的行为如何? 给出几个例子.

(b) 如果 $f'(x_0) = 0$ 而 $f''(x_0) \neq 0$ 呢?

(c) 如果 $f'(x_0) = f''(x_0) = 0$ 而 $f'''(x_0) \neq 0$ 呢?

14. 考虑一阶非自治方程 $x' = p(t)x$, 其中 $p(t)$ 是周期为 T 的可微函数. 证明该方程的所有解都是以 T 为周期的当且仅当

$$\int_0^T p(s)\mathrm{d}s = 0.$$

15. 考虑微分方程 $x' = f(t,x)$, 其中 $f(t,x)$ 对 t 和 x 连续可微. 假设对所有的 t 都有

$$f(t+T,x) = f(t,x).$$

再假设存在常数 p,q, 使得对所有的 t 都有

$$f(t,p) > 0, \quad f(t,q) < 0.$$

证明该方程有一个满足 $p < x(0) < q$ 的周期解 $x(t)$.

16. 考虑微分方程 $x' = x^2 - 1 - \cos(t)$. 讨论该方程周期解的存在性.

第 2 章　平面线性系统

在本章我们开始研究**微分方程组系统**. 一个微分方程组系统是指如下的互相关联的 n 个微分方程

$$\begin{aligned} x_1' &= f_1(t, x_1, x_2, \cdots, x_n), \\ x_2' &= f_2(t, x_1, x_2, \cdots, x_n), \\ &\vdots \\ x_n' &= f_n(t, x_1, x_2, \cdots, x_n). \end{aligned}$$

这里 f_j 是 $n+1$ 个变量 $x_1, x_2, \cdots, x_n$ 和 t 的实值函数. 除非另外声明, 我们总假设 f_j 是 C^∞ 函数. 这意味着 f_j 的各阶偏导数都存在且连续.

为简化记号, 我们使用向量符号

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

我们经常将向量 $\boldsymbol{X}$ 写成 $(x_1, \cdots, x_n)$ 以节省版面空间.

这样我们的系统就可以写成如下的简洁形式：

$$\boldsymbol{X}' = F(t, \boldsymbol{X}),$$

其中

$$F(t, \boldsymbol{X}) = \begin{pmatrix} f_1(t, x_1, x_2, \cdots, x_n) \\ \vdots \\ f_n(t, x_1, x_2, \cdots, x_n) \end{pmatrix}.$$

该系统的一个解就是满足方程的形如 $\boldsymbol{X}(t) = (x_1(t), \cdots, x_n(t))$ 的函数, 即

$$\boldsymbol{X}'(t) = F(t, \boldsymbol{X}(t)),$$

其中 $\boldsymbol{X}'(t) = (x_1'(t), \cdots, x_n'(t))$. 当然, 现在我们还不能保证存在这样的一个解, 但我们将在 2.7 节讨论这个复杂的问题.

如果所有的 f_j 都不依赖 t, 上述方程组系统就称为**自治的**, 此时系统变成 $\boldsymbol{X}' = F(\boldsymbol{X})$. 在本书后面的大部分, 我们将主要关心自治系统.

同一阶微分方程一样, 满足 $F(\boldsymbol{X}_0)=0$ 的向量 $\boldsymbol{X}_0$ 称为系统的一个**平衡点**, 而该平衡点则对应于系统的一个常数解 $\boldsymbol{X}(t)\equiv\boldsymbol{X}_0$.

下面对本书后面采用的符号作一些说明：小写字母, 如 x,y,x_1,x_2,t 等, 将用来代表实变量; 而实值函数则用 $f(x,y)$ 或 $f_1(x_1,\cdots,x_n,t)$ 这样的小写字母表示; 大写字母则用来表示向量, 如 $\boldsymbol{X}=(x_1,\cdots,x_n)$, 或向量值函数, 如

$$\boldsymbol{F}(x,y)=(f(x,y),g(x,y))$$

或

$$\boldsymbol{H}(x_1,\cdots,x_n)=\begin{pmatrix}h_1(x_1,\cdots,x_n)\\ \vdots\\ h_n(x_1,\cdots,x_n)\end{pmatrix}.$$

我们将用 $\mathbb{R}^n$ 代表 n 维欧氏空间, 于是 $\mathbb{R}^n$ 由所有形如 $\boldsymbol{X}=(x_1,\cdots,x_n)$ 的向量构成.

2.1 二阶微分方程

在科学和工程中出现的许多最重要的微分方程都是二阶微分方程. 这些方程具有如下的形式：

$$x''=f(t,x,x').$$

这些二阶方程的重要例子中包括牛顿方程

$$mx''=f(x),$$

电子工程中的 RLC 电路方程

$$LCx''+RCx'+x=\nu(t),$$

以及大多数初等微分方程课程的支柱 —— 受迫调和振子

$$mx''+bx'+kx=f(t).$$

随着课程的展开, 我们将深入讨论这些方程以及它们之间的关系. 但是, 首先注意到, 如果引入第二个变量 $y=x'$, 这些方程都只是一般二维微分方程系统的一类特例.

例如, 考虑如下的常系数二阶方程

$$x''+ax'+bx=0.$$

如果令 $y=x'$, 则可将上述方程改写为以下的一阶方程组

$$x'=y$$
$$y'=-bx-ay.$$

任何二阶方程都可以通过这种方式改写成一个一阶方程组. 因而在本书的后面, 我们将主要讨论方程组.

2.2 平面系统

在本章的余下部分, 我们将讨论平面 $\mathbb{R}^2$ 上的自治系统. 将它们写成如下形式

$$x'=f(x,y),$$
$$y'=g(x,y),$$

这样我们就扔掉了出现在函数和变量中令人讨厌的下标. 同前面一样, 我们将上述方程简写成 $\boldsymbol{X}'=F(\boldsymbol{X})$, 其中 $\boldsymbol{X}=(x,y)$, $\boldsymbol{F}(\boldsymbol{X})=\boldsymbol{F}(x,y)=(f(x,y),g(x,y))$.

与第 1 章中的斜率场类似, 我们将上述方程的右端看成是 $\mathbb{R}^2$ 上的**向量场**, 即我们认为 $\boldsymbol{F}(x,y)$ 代表 $\mathbb{R}^2$ 上的一个向量, 其 x 分量和 y 分量分别为 $f(x,y)$ 和 $g(x,y)$, 而且其基点为 (x,y). 例如, 与系统

$$x'=y$$
$$y'=-x$$

相关联的向量场如图 2.1 所示. 注意到, 此时许多向量相互重叠, 这使得图形很难分辨. 为此, 我们通常作**方向场**来替代. 所谓方向场就是向量场的归一化, 即不考虑向量的长度而只考虑其方向.

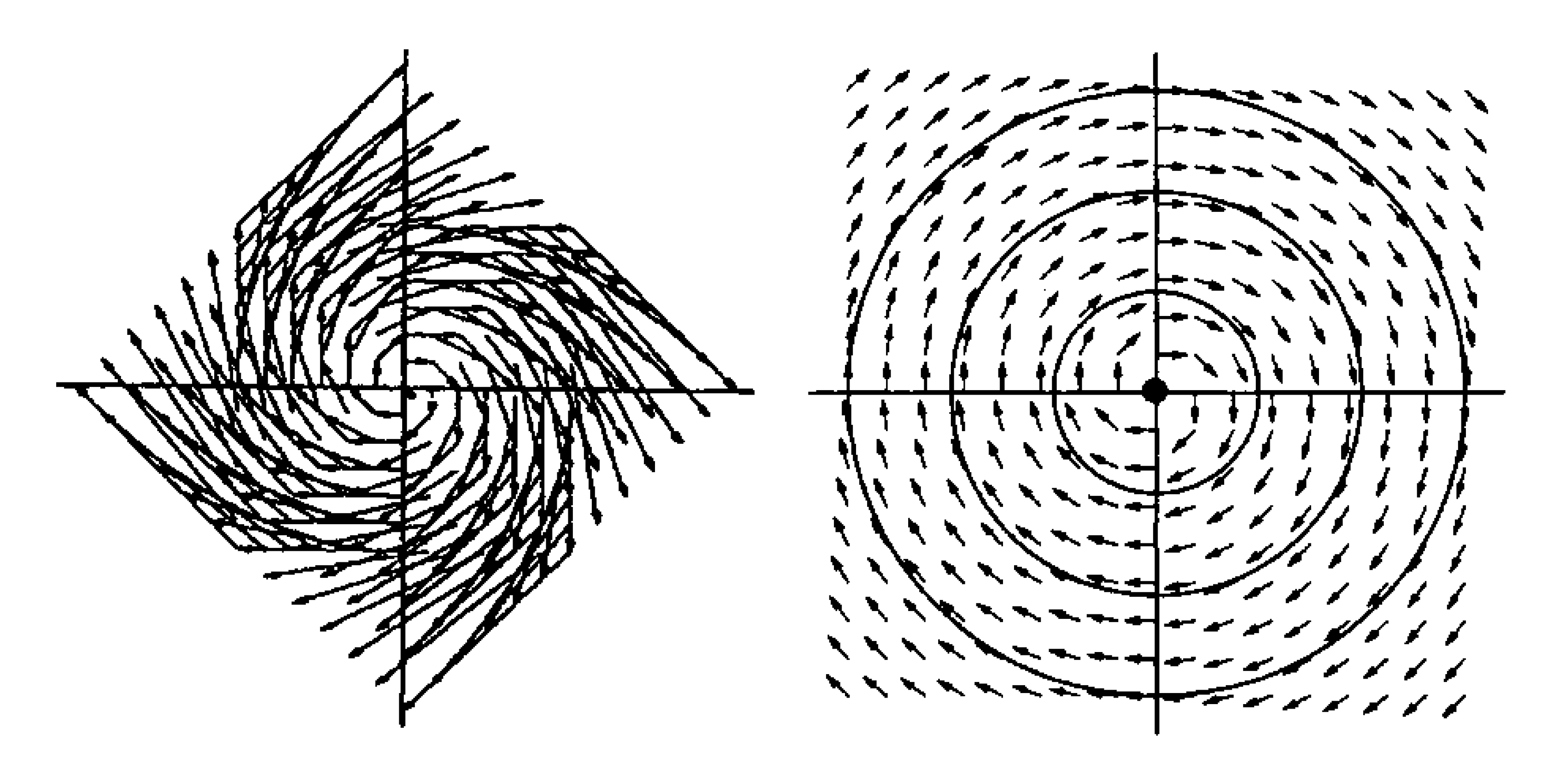


图 2.1 系统 $x'=y, y'=-x$ 的向量场、方向场和一些解

现在, 该系统的一个解可以看成是平面上一条形如 $(x(t),y(t))$ 的参数曲线, 其中对每一个 t, 在点 $(x(t),y(t))$ 处的切向量就是 $\boldsymbol{F}(x(t),y(t))$, 即, 解曲线 $(x(t),y(t))$ 总是以在基点 $(x(t),y(t))$ 处切于给定向量 $\boldsymbol{F}(x(t),y(t))$ 的方式在平面上缠绕.

例 任给 $a \in \mathbb{R}$; 曲线

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} a\sin t \\ a\cos t \end{pmatrix}$$

都是系统

$$\begin{aligned} x' &= y, \\ y' &= -x, \end{aligned}$$

的一个解, 理由如下:

$$\begin{aligned} x'(t) &= a\cos t = y(t), \\ y'(t) &= -a\sin t = -x(t). \end{aligned}$$

这些曲线是平面上半径为 $|a|$ 的圆周, 当 t 增加时, 曲线沿顺时针方向旋转. 当 $a = 0$ 时, 解为常数函数 $x(t) \equiv 0 \equiv y(t)$. ■

注意, 通过引入第二个变量 $y = x'$, 二阶微分方程 $x'' = -x$ 与上述例子等价. 这是一个二阶**线性**微分方程的例子, 其更一般的形式是

$$a(t)x'' + b(t)x' + c(t)x = f(t).$$

其中一个重要的情形是所谓的线性**常系数**方程

$$ax'' + bx' + cx = f(t),$$

它可以改写成如下形式

$$\begin{aligned} x' &= y \\ y' &= -\frac{c}{a}x - \frac{b}{a}y + \frac{f(t)}{a}. \end{aligned}$$

一个更特殊的情形是**齐次**方程, 此时 $f(t) \equiv 0$.

例 一个最简单也是最重要的二阶线性常系数微分方程就是**调和振子**的方程. 该方程描述了与一个弹簧相连的质点运动. 弹簧的另一端固定在一垂直的墙上, 而容许质点沿水平方向滑动. 用 x 代表质点离开其自然平衡位置的位移 ($x > 0$ 表示弹簧被拉伸, 而 $x < 0$ 则表示弹簧被压缩). 于是 $x'(t)$ 就是运动质点的速度, $x''(t)$ 为加速度. 此时, 弹簧有一个正比于 $x(t)$ 的弹性回复力, 此外还有一个与运动方向相反、正比于 $x'(t)$ 的摩擦力. 该系统有三个参数: m 为振子的质量, $b \geqslant 0$ 为**阻尼系数**, $k > 0$ 为**弹性系数**. 牛顿定律断言振子上的作用力等于其质量乘以加速度. 于是, 带阻尼的调和振子的微分方程为:

$$mx'' + bx' + kx = 0.$$

当 $b=0$ 时, 振子称为**无阻尼的**, 否则称为**带阻尼的调和振子**. 这是一个二阶线性常系数齐次微分方程的例子. 作为一个系统, 调和振子方程成为

$$x' = y$$

$$y' = -\frac{k}{m}x - \frac{b}{m}y.$$

更一般地, 质点弹簧系统还会受到一外力作用, 如将垂直墙面前后周期移动. 这样的外力通常只依赖于时间, 而与位置无关, 于是我们得到下面的受迫调和振子系统

$$mx'' + bx' + kx = f(t),$$

其中 $f(t)$ 代表外力. 这是一个非自治的二阶线性方程. ■

2.3 代数预备知识

在进一步讨论微分方程系统之前, 我们需要先提一下相关代数方程组系统的一些初等事实. 我们将常常遇到如下的方程组:

$$ax + by = \alpha,$$
$$cx + dy = \beta,$$

其中 a, b, c, d 以及 α, β 都是事先给定的. 上述方程可用矩阵形式写成

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

用 $\boldsymbol{A}$ 表示 2×2 的系数矩阵

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

当假设该方程组有解时, 它的解是很容易求的. 事实上, 这样的方程组有唯一解的充要条件是 $\boldsymbol{A}$ 的**行列式**非零. 回忆行列式的定义

$$\det \boldsymbol{A} = ad - bc.$$

当 $\det \boldsymbol{A} = 0$ 时, 可能有解也可能无解, 但如果有解的话, 则事实上有无穷多解.

如果 $\det \boldsymbol{A} = 0$, 在 $\alpha = \beta = 0$ 的特殊情形,

$$\boldsymbol{A} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

总有无穷多解. 事实上, 如果 $\boldsymbol{A}$ 中的系数 a 非零, 则有 $x=-(b/a)y$, 从而

$$-c\left(\frac{b}{a}\right)y+dy=0.$$

于是 $(ad-bc)y=0$. 由于 $\det \boldsymbol{A}=0$, 方程的解为 $(-(b/a)y,y)$, 其中 y 是任意的. 这意味着, 任一解都位于平面上过原点的一条直线上. 只要 $\boldsymbol{A}$ 中的元素有一个非零, 都会有类似的解直线. 我们不用太担心 $\boldsymbol{A}$ 中的元素都为零的情形, 事实上, 我们将完全不关心这种情况.

设 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 为平面上的两个向量. 称两个非零向量 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 为**线性无关的**, 如果 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 不在过原点的同一条直线上. 向量 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 称为**线性相关的**, 如果 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 中有一个为零向量, 或它们在过原点的同一条直线上.

平面上两个向量线性无关的一个几何判据是它们的指向既不相同也不相反. 即, 两个非零 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 是线性无关的当且仅当对任意的实数 λ, $\boldsymbol{V}\neq\lambda\boldsymbol{W}$. 线性无关一个等价的代数判据为：

命题 假设向量 $\boldsymbol{V}=(v_1,v_2)$, $\boldsymbol{W}=(w_1,w_2)$. 则 $\boldsymbol{V}$ 与 $\boldsymbol{W}$ 是线性无关的当且仅当

$$\det\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}\neq 0.$$

证明留作练习, 见本章的习题 11. ■

只要有了一对线性无关的向量 $\boldsymbol{V}$ 和 $\boldsymbol{W}$, 则任意给定的向量 $\boldsymbol{Z}\in\mathbb{R}^2$ 都可以唯一地写成 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 的**线性组合**, 也就是说可以找到实数对 α 和 β 使得

$$\boldsymbol{Z}=\alpha\boldsymbol{V}+\beta\boldsymbol{W}.$$

并且, 这样的 α 和 β 是唯一的. 为此, 假设 $\boldsymbol{Z}=(z_1,z_2)$. 于是就要求解方程组

$$\begin{aligned} z_1&=\alpha v_1+\beta w_1,\\ z_2&=\alpha v_2+\beta w_2, \end{aligned}$$

这里 v_i,w_i 和 z_i 是已知的. 但是因为

$$\det\begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix}\neq 0,$$

上述方程组有唯一解 (α,β). 我们称线性无关的向量 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 为 $\mathbb{R}^2$ 的一个**基**. 相对于 $\boldsymbol{V}$ 和 $\boldsymbol{W}$, 任一向量 $\boldsymbol{Z}$ 都有唯一的"坐标", 即满足 $\boldsymbol{Z}=\alpha\boldsymbol{V}+\beta\boldsymbol{W}$ 二元数组 (α,β).

例 显然单位向量 $\boldsymbol{E}_1=(1,0)$ 和 $\boldsymbol{E}_2=(0,1)$ 构成了 $\mathbb{R}^2$ 的一个基, 这个基称为 $\mathbb{R}^2$ 的**标准基**. $\boldsymbol{Z}$ 在标准基下的坐标就是 $\boldsymbol{Z}$ 的 "通常" 笛卡儿坐标 (x,y). ■

例 向量 $\boldsymbol{V}_1=(1,1)$ 和 $\boldsymbol{V}_1=(-1,1)$ 也构成了 $\mathbb{R}^2$ 的一个基. 在这个基下, $\boldsymbol{E}_1$ 的坐标是 $(1/2,-1/2)$, 而 $\boldsymbol{E}_2$ 的坐标是 $(1/2,1/2)$, 理由如下:

$$\begin{pmatrix}1\\0\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}-\frac{1}{2}\begin{pmatrix}-1\\1\end{pmatrix}$$

$$\begin{pmatrix}0\\1\end{pmatrix}=\frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}+\frac{1}{2}\begin{pmatrix}-1\\1\end{pmatrix}$$

这种 "坐标变换" 在后面将会变得重要. ■

例 向量 $\boldsymbol{V}_1=(1,1)$ 和 $\boldsymbol{V}_1=(-1,-1)$ 不能构成 $\mathbb{R}^2$ 的一个基, 因为它们是共线的. 这两个向量的任一线性组合都具有如下形式:

$$\alpha\boldsymbol{V}_1+\beta\boldsymbol{V}_2=\begin{pmatrix}\alpha-\beta\\\alpha-\beta\end{pmatrix},$$

它们都位于过原点、$\boldsymbol{V}_1$ 及 $\boldsymbol{V}_2$ 的直线上. ■

2.4 平面线性系统

我们现在将注意力集中到最重要的一类平面微分方程系统, 也即线性系统. 在自治情形, 这些系统有如下简单的形式

$$\begin{aligned}x'&=ax+by,\\y'&=cx+dy,\end{aligned}$$

这里 a,b,c,d 为常数. 记

$$\boldsymbol{A}=\begin{pmatrix}a&b\\c&d\end{pmatrix}$$

为系数矩阵, 此时系统可以简写成

$$\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}.$$

注意到原点总是线性系统的一个平衡点. 为了寻找其他的平衡点, 我们需要求解代数方程组的线性系统

$$\begin{aligned}ax+by&=0,\\cx+dy&=0.\end{aligned}$$

这个系统有非零解当且仅当 $\det \boldsymbol{A}=0$. 在前面我们已经知道, 当 $\det \boldsymbol{A}=0$ 时, 则有一条通过原点的直线, 上面的每一个点都是平衡点. 于是我们得到:

命题

(1) 当 $\det \boldsymbol{A} \neq 0$ 时, 平面线性系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 有唯一的平衡点 $(0,0)$;

(2) 当 $\det \boldsymbol{A}=0$(并且 $\boldsymbol{A}$ 不是 0 矩阵) 时, 平面线性系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的平衡点由一条直线构成. ■

2.5 特征值和特征向量

现在我们来研究如何寻找线性系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的非平衡点解. 此时, 一个关键的观察是: 假设 $\boldsymbol{V}_0$ 为一非零向量, 满足 $\boldsymbol{A}\boldsymbol{V}_0=\lambda\boldsymbol{V}_0(\lambda\in\mathbb{R})$, 则函数

$$\boldsymbol{X}(t)=\mathrm{e}^{\lambda t}\boldsymbol{V}_0$$

为系统的一个解. 为此, 我们作如下计算:

$$\begin{aligned}\boldsymbol{X}'(t)&=\lambda\mathrm{e}^{\lambda t}\boldsymbol{V}_0\\&=\mathrm{e}^{\lambda t}(\lambda\boldsymbol{V}_0)\\&=\mathrm{e}^{\lambda t}(\boldsymbol{A}\boldsymbol{V}_0)\\&=\boldsymbol{A}(\mathrm{e}^{\lambda t}\boldsymbol{V}_0)\\&=\boldsymbol{A}\boldsymbol{X}(t).\end{aligned}$$

这表明 $\boldsymbol{X}(t)$ 确实为系统的一个解. 这样的向量 $\boldsymbol{V}_0$ 以及与之相关的标量有下面的名称:

定义 一个非零向量 $\boldsymbol{V}_0$ 称为 $\boldsymbol{A}$ 的一个**特征向量**, 如果对某个实数 λ, $\boldsymbol{A}\boldsymbol{V}_0=\lambda\boldsymbol{V}_0$, 常数 λ 则称为 $\boldsymbol{A}$ 的一个**特征值**. ■

刚才我们已经看到, 特征值、特征向量和微分方程系统的解之间有一个重要的关系:

定理 假设向量 $\boldsymbol{V}_0$ 为矩阵 $\boldsymbol{A}$ 属于 λ 的一个特征向量, 则函数 $\boldsymbol{X}(t)=\mathrm{e}^{\lambda t}\boldsymbol{V}_0$ 为系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的一个解. ■

如果 $\boldsymbol{V}_0$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的一个特征向量, 则 $\boldsymbol{V}_0$ 乘以任何一个非零的标量后依然是 $\boldsymbol{A}$ 的一个特征向量. 事实上, 如果 $\boldsymbol{A}\boldsymbol{V}_0=\lambda\boldsymbol{V}_0$, 则对任何非零常数 α, 我们有

$$\boldsymbol{A}(\alpha\boldsymbol{V}_0)=\alpha\boldsymbol{A}\boldsymbol{V}_0=\lambda(\alpha\boldsymbol{V}_0).$$

例 考虑矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 3\\1 & -1\end{pmatrix}.$$

则 $\boldsymbol{A}$ 有特征向量 $\boldsymbol{V}_0=(3,1)$, 相应的特征值为 $\lambda=2$, 理由如下:

$$\begin{pmatrix}1 & 3\\ 1 & -1\end{pmatrix}\begin{pmatrix}3\\ 1\end{pmatrix}=\begin{pmatrix}6\\ 2\end{pmatrix}=2\begin{pmatrix}3\\ 1\end{pmatrix}.$$

■

类似地, $\boldsymbol{V}_1=(1,-1)$ 是属于特征值 $\lambda=-2$ 的特征向量.

于是, 对于系统

$$\boldsymbol{X}'=\begin{pmatrix}1 & 3\\ 1 & -1\end{pmatrix}\boldsymbol{X},$$

我们已经知道了如下三个解: 位于原点的平衡解以及

$$\boldsymbol{X}_1(t)=\mathrm{e}^{2t}\begin{pmatrix}3\\ 1\end{pmatrix},\quad \boldsymbol{X}_2(t)=\mathrm{e}^{-2t}\begin{pmatrix}1\\ -1\end{pmatrix}.$$

后面我们将看到, 可以利用这些解立刻生成系统的**所有的**解, 但现在我们还是来看看如何寻找特征向量和特征值.

为了得到一个特征向量 $\boldsymbol{V}=(x,y)$, 我们需要寻找方程

$$\boldsymbol{A}\begin{pmatrix}x\\ y\end{pmatrix}=\lambda\begin{pmatrix}x\\ y\end{pmatrix}$$

的一个非零解 (x,y). 在这个方程组系统中有三个未知量: $\boldsymbol{V}$ 的两个分量以及 λ. 记 $\boldsymbol{I}$ 为 2×2 的单位矩阵

$$\boldsymbol{I}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

则上述方程可以改写为

$$(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}=\boldsymbol{0},$$

其中 $\boldsymbol{0}$ 代表向量 $(0,0)$.

$\boldsymbol{A}-\lambda\boldsymbol{I}$ 也是一个 2×2 矩阵, 其元素含有变量 λ. 在前面我们已经知道线性方程组有非零解当且仅当 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0$. 但该方程只是一个关于 λ 的二次方程, 很容易求得它的根. 方程 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})=0$ 在随后将会频繁出现, 它称为**特征方程**. 作为 λ 的函数, $\det(\boldsymbol{A}-\lambda\boldsymbol{I})$ 称为**特征多项式**. 因而为了寻找特征向量, 首先需要找到特征方程的根, 也就是特征值. 然后利用这些特征值再来寻找相应的特征向量.

例 仍然回到矩阵

$$\boldsymbol{A}=\begin{pmatrix}1 & 3\\ 1 & -1\end{pmatrix}.$$

我们有

$$A - \lambda I = \begin{pmatrix} 1-\lambda & 3 \\ 1 & -1-\lambda \end{pmatrix}.$$

从而特征方程为

$$\det(A - \lambda I) = (1-\lambda)(-1-\lambda) - 3 = 0,$$

化简后得到

$$\lambda^2 - 4 = 0,$$

由此可解得特征值 $\lambda = \pm 2$. 对于 $\lambda = 2$, 下面我们来解方程

$$(A - 2I)\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

写成分量形式后, 我们得到方程组系统

$$\begin{aligned}(1-2)x + 3y &= 0,\\ x + (-1-2)y &= 0,\end{aligned}$$

也就是 $-x + 3y = 0$(因为第二个方程是多余的). 从而任何形如 $(3y, y)(y \neq 0)$ 的向量都是 $\lambda = 2$ 的一个特征向量. 用同样的方式, 可以得到, 任何形如 $(y, -y)(y \neq 0)$ 的向量都是 $\lambda = -2$ 的一个特征向量. ■

当然, 聪明的读者将注意到, 特征值、特征向量和微分方程的解之间的故事不仅仅只是我们刚刚所描述的那些. 例如, 特征方程的根可能是复的, 也有可能是相同的实数. 我们将简短地处理这些情形, 但在此之前, 我们还是先看看如何求解线性系统的问题.

2.6 求解线性系统

在上节我们已经看到, 如果我们可以找到特征方程的两个不同的实根 λ_1 和 λ_2, 则我们就可以得到微分方程系统的一对解 $X_i(t) = e^{\lambda_i t}V_i$, 这里 V_i 是属于 λ_i 的一个特征向量. 注意这每一个解都是**直线解**. 事实上, 我们有 $X_i(0) = V_i$, 这是平面上的一个非零点. 对每个 t, $e^{\lambda_i t}V_i$ 无非就是 V_i 乘上一个标量, 因而它们都位于从原点出发经过 V_i 的射线上. 而且, 如果 $\lambda_i > 0$, 则有

$$\lim_{t\to\infty} |X_i(t)| = \infty$$

以及

$$\lim_{t\to-\infty} X_i(t) = (0, 0).$$

当 t 增加时, 沿着通过 $\boldsymbol{V}_i$ 的射线, 解 $\boldsymbol{X}_i(t)$ 单调增加到 ∞, 而当时间反向时, $\boldsymbol{X}_i(t)$ 则沿着这条射线趋于原点. 当 $\lambda_i < 0$ 时, 情形则正好相反, 而当 $\lambda_i = 0$ 时, $\boldsymbol{X}_i(t)$ 为常数解, 即对所有的 t, $\boldsymbol{X}_i(t) = \boldsymbol{V}_i$.

已知这两个特解后如何找到系统的全部解呢? 答案是简要的. 假设有两个不同的实特征值 λ_1 和 λ_2 以及相应的两个特征向量 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$. 容易验证 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 是线性无关的 (见习题 14). 于是 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 构成了 $\mathbb{R}^2$ 的一个基, 从而任给点 $\boldsymbol{Z}_0 \in \mathbb{R}^2$, 可以找到唯一的一对实数 α 和 β 满足

$$\alpha \boldsymbol{V}_1 + \beta \boldsymbol{V}_2 = \boldsymbol{Z}_0.$$

现在考虑函数 $\boldsymbol{Z}(t) = \alpha \boldsymbol{X}_1(t) + \beta \boldsymbol{X}_2(t)$, 这里 $\boldsymbol{X}_i(t)$ 就是刚才的直线解. 我们断言 $\boldsymbol{Z}(t)$ 是 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的一个解. 为此我们来计算

$$\begin{aligned}
\boldsymbol{Z}'(t) &= \alpha \boldsymbol{X}_1'(t) + \beta \boldsymbol{X}_2'(t) \\
&= \alpha \boldsymbol{A}\boldsymbol{X}_1(t) + \beta \boldsymbol{A}\boldsymbol{X}_2(t) \\
&= \boldsymbol{A}(\alpha \boldsymbol{X}_1(t) + \beta \boldsymbol{X}_2(t)) \\
&= \boldsymbol{A}\boldsymbol{Z}(t).
\end{aligned}$$

其中的倒数第二步来自于矩阵乘法的线性性 (见习题 13). 这样我们就说明了 $\boldsymbol{Z}'(t) = \boldsymbol{A}\boldsymbol{Z}(t)$, 因而 $\boldsymbol{Z}(t)$ 为一个解. 而且, $\boldsymbol{Z}(t)$ 是满足 $\boldsymbol{Z}(0) = \boldsymbol{Z}_0$ 的解. 最后, 我们断言 $\boldsymbol{Z}(t)$ 是 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 满足 $\boldsymbol{Z}(0) = \boldsymbol{Z}_0$ 的唯一解. 像在第 1 章所做的一样, 假设 $\boldsymbol{Y}(t)$ 是另一个满足 $\boldsymbol{Y}(0) = \boldsymbol{Z}_0$ 的解. 于是我们就有

$$\boldsymbol{Y}(t) = \zeta(t)\boldsymbol{V}_1 + \mu(t)\boldsymbol{V}_2,$$

其中 $\zeta(0) = \alpha$, $\mu(0) = \beta$. 于是,

$$\boldsymbol{A}\boldsymbol{Y}(t) = \boldsymbol{Y}'(t) = \zeta'(t)\boldsymbol{V}_1 + \mu'(t)\boldsymbol{V}_2.$$

另一方面,

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{Y}(t) &= \zeta(t)\boldsymbol{A}\boldsymbol{V}_1 + \mu(t)\boldsymbol{A}\boldsymbol{V}_2 \\
&= \lambda_1\zeta(t)\boldsymbol{V}_1 + \lambda_2\mu(t)\boldsymbol{V}_2.
\end{aligned}$$

这样我们就得到

$$\begin{aligned}
\zeta'(t) &= \lambda_1\zeta(t), \\
\mu'(t) &= \lambda_2\mu(t),
\end{aligned}$$

以及 $\zeta(0) = \alpha$, $\mu(0) = \beta$. 从第 1 章的结论立得

$$\zeta(t) = \alpha \mathrm{e}^{\lambda_1 t}, \quad \mu(t) = \beta \mathrm{e}^{\lambda_2 t},$$

从而 $\boldsymbol{Y}(t)$ 就等于 $\boldsymbol{Z}(t)$.

作为一个推论, 我们就得到了, 对于任意给定的 $\boldsymbol{Z}_0 \in \mathbb{R}^2$, 系统 $\boldsymbol{X}' = \boldsymbol{AX}$ 有唯一的解满足 $\boldsymbol{X}(0) = \boldsymbol{Z}_0$. 所有这些解的全体称为 $\boldsymbol{X}' = \boldsymbol{AX}$ 的**通解**. 即, $\boldsymbol{X}' = \boldsymbol{AX}$ 的通解是指对每一 $\boldsymbol{Z}_0 \in \mathbb{R}^2$, 初值问题 $\boldsymbol{X}(0) = \boldsymbol{Z}_0$ 的唯一解构成的全体.

我们已经证明:

定理 假设 $\boldsymbol{A}$ 有一对实特征值 $\lambda_1 \neq \lambda_2$, 相应的特征向量为 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$. 则线性系统 $\boldsymbol{X}' = \boldsymbol{AX}$ 的通解为

$$\boldsymbol{X}(t) = \alpha \mathrm{e}^{\lambda_1 t}\boldsymbol{V}_1 + \beta \mathrm{e}^{\lambda_2 t}\boldsymbol{V}_2.$$ ■

例 考虑二阶微分方程

$$x'' + 3x' + 2x = 0.$$

这是早先讨论的带阻尼的调和振子的一个特例, 其中质量为 1, 弹性系数为 2, 阻尼系数为 3. 作为一个系统, 这个方程可以改写为

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 \\ -2 & -3 \end{pmatrix} \boldsymbol{X} = \boldsymbol{AX}.$$

其特征方程为

$$\lambda^2 + 3\lambda + 2 = (\lambda + 2)(\lambda + 1) = 0,$$

从而系统的特征值为 -1 和 -2. 属于特征值 -1 的特征向量可以通过求解方程

$$(\boldsymbol{A} + \boldsymbol{I}) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

得到. 上述方程的分量形式为

$$\begin{aligned} x + y &= 0, \\ -2x - 2y &= 0. \end{aligned}$$

从而, $(1, -1)$ 是属于特征值 -1 的一个特征向量. 类似的计算可得 $(1, -2)$ 是属于特征值 -2 的一个特征向量. 易见这两个特征向量是线性无关的. 于是根据前面的定理, 该系统的通解为

$$\boldsymbol{X}(t) = \alpha \mathrm{e}^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \beta \mathrm{e}^{-2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

即, 质点的位置由解的第一个分量给出

$$x(t) = \alpha \mathrm{e}^{-t} + \beta \mathrm{e}^{-2t},$$

而质点的速度则由第二个分量给出

$$y(t) = x'(t) = -\alpha e^{-t} - 2\beta e^{-2t}. \qquad \blacksquare$$

2.7　线性叠加原理

上一节讨论的定理是一般 n 维线性系统基本定理一个很特殊的情形, 这个一般的定理将在 6.1 节中证明. 这个结论在二维情形可以如下描述：若 $\boldsymbol{X}' = \boldsymbol{AX}$ 是一平面线性系统, $\boldsymbol{Y}_1(t)$ 和 $\boldsymbol{Y}_2(t)$ 是它的两个解, 则函数 $\alpha\boldsymbol{Y}_1(t) + \beta\boldsymbol{Y}_2(t)$ 也是该系统的一个解. 证明这个结论并没有用到特征值是实的且是不同的. 这个事实称为**线性叠加原理**. 更重要的是, 如果初值条件 $\boldsymbol{Y}_1(0)$ 和 $\boldsymbol{Y}_2(0)$ 是线性无关的向量, 则它们构成了 $\mathbb{R}^2$ 的一个基, 于是任给向量 $\boldsymbol{X}_0 \in \mathbb{R}^2$, 可以确定常数 α 和 β 使得 $\boldsymbol{X}_0 = \alpha\boldsymbol{Y}_1(0) + \beta\boldsymbol{Y}_2(0)$, 于是线性叠加原理告诉我们, 满足初值条件 $\boldsymbol{X}(0) = \boldsymbol{X}_0$ 的解就是 $\boldsymbol{X}(t) = \alpha\boldsymbol{Y}_1(t) + \beta\boldsymbol{Y}_2(t)$. 从而我们就得到了系统任一给定初值问题的一个解. 第 6 章关于线性系统的存在唯一性定理将说明这样的解是唯一的. 这一重要的结论可以总结成：

定理　设 $\boldsymbol{X}' = \boldsymbol{AX}$ 为一平面线性系统, $\boldsymbol{Y}_1(t)$ 和 $\boldsymbol{Y}_2(t)$ 是它的解, 而且向量 $\boldsymbol{Y}_1(0)$ 和 $\boldsymbol{Y}_2(0)$ 是线性无关的. 则

$$\boldsymbol{X}(t) = \alpha\boldsymbol{Y}_1(t) + \beta\boldsymbol{Y}_2(t)$$

是系统满足初值条件 $\boldsymbol{X}_0 = \alpha\boldsymbol{Y}_1(0) + \beta\boldsymbol{Y}_2(0)$ 的唯一解. $\blacksquare$

习　　题

1. 找出下列 2×2 矩阵的特征值和特征向量：

   (a) $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$　(b) $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$　(c) $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$　(d) $\begin{pmatrix} 1 & 3 \\ \sqrt{2} & 3\sqrt{2} \end{pmatrix}$

2. 找出下列线性系统的通解：

   (a) $\boldsymbol{X}' = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}\boldsymbol{X}$　(b) $\boldsymbol{X}' = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}\boldsymbol{X}$

   (c) $\boldsymbol{X}' = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}\boldsymbol{X}$　(d) $\boldsymbol{X}' = \begin{pmatrix} 1 & 2 \\ 3 & -3 \end{pmatrix}\boldsymbol{X}$

3. 试将图 2.2 中的四个方向场与上一题中的四个系统一一对应.

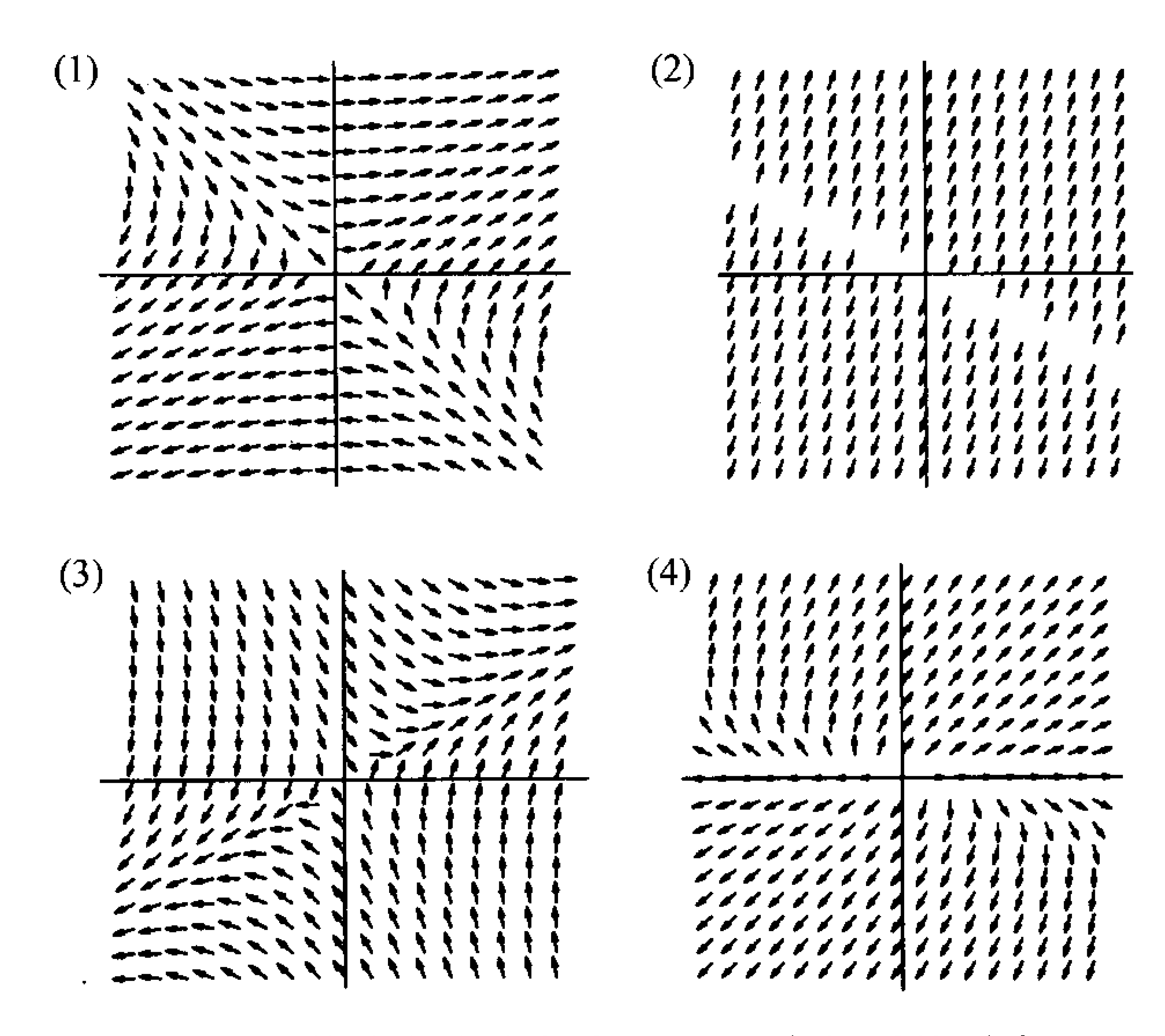


图 2.2　将这些方向场与习题 2 中的系统对应

4. 找出系统

$$X' = \begin{pmatrix} a & b \\ c & a \end{pmatrix} X$$

的通解, 其中 $bc > 0$.

5. 找出系统

$$X' = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} X$$

的通解.

6. 考虑调和振子系统 $x'' + bx' + kx = 0$, 找出所有 b, k 的取值, 使得该系统具有实的且不同的特征值. 并求此时系统的通解. 找出满足初值条件 $(0,1)$ 的解. 并描述此时质点的运动.

7. 考虑 2×2 矩阵

$$A = \begin{pmatrix} a & 1 \\ 0 & 1 \end{pmatrix}.$$

找出参数 a 的取值 a_0, 使得 A 有实的重特征值. 当 a 趋于 a_0 时, 矩阵的特征向量将会如何变化?

8. 描述特征值为 0 和 1 的所有可能的 2×2 矩阵.

9. 给出一个线性系统的例子, 使得对任一常数 α, $(\mathrm{e}^{-t}, \alpha)$ 都是系统的解, 并作该系统方向场的简图. 系统的通解是什么?

10. 给出一个微分方程系统的例子, 使得 $(t, 1)$ 是系统的一个解, 并作该系统的方向场的简图. 系统的通解是什么?
11. 证明向量 $\boldsymbol{V} = (v_1, v_2)$ 和 $\boldsymbol{W} = (w_1, w_2)$ 线性无关的充要条件是

$$\det \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \neq 0.$$

12. 证明: 如果 λ, μ 是 2×2 矩阵 $\boldsymbol{A}$ 的实特征值, 则矩阵 $\boldsymbol{A} - \lambda \boldsymbol{I}$ 的非零列都是 μ 的特征向量.
13. 设 $\boldsymbol{A}$ 为 2×2 矩阵, $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 都是 $\mathbb{R}^2$ 中的向量, 试证明 $\boldsymbol{A}(\alpha \boldsymbol{V}_1 + \beta \boldsymbol{V}_2) = \alpha \boldsymbol{A} \boldsymbol{V}_1 + \beta \boldsymbol{A} \boldsymbol{V}_2$.
14. 证明一个 2×2 矩阵的两个不同实特征值的特征向量是线性无关的.

第 3 章　平面系统的相图

有了上一章的线性叠加原理后，我们现在来计算任一平面系统的通解．粗看，似乎有无穷多不同的情形要讨论，但我们将看到，最简形式的几个例子就几乎涵盖了我们在高维情形将要遇到的所有解的类型．

3.1　不同实特征值

考虑系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$，假设 $\boldsymbol{A}$ 有两个实特征值 $\lambda_1 < \lambda_2$．先暂时假设 $\lambda_i \neq 0$，此时有如下三种情形：

(1) $\lambda_1 < 0 < \lambda_2$；　(2) $\lambda_1 < \lambda_2 < 0$；　(3) $0 < \lambda_1 < \lambda_2$．

我们先对每种情形给出一个典型例子，随后我们将看到任何属于这三类的系统都可以类似地处理．

例　(**鞍点**) 首先考虑简单的系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$，其中

$$\boldsymbol{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

并且 $\lambda_1 < 0 < \lambda_2$．该系统可以分解成两个不相关联的一阶方程

$$\begin{aligned} x' &= \lambda_1 x \\ y' &= \lambda_2 y. \end{aligned}$$

我们已经知道如何求解这两个方程，但为了与后面相对照，我们还是先找特征值和特征向量．系统的特征方程是

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = 0,$$

于是 λ_1 和 λ_2 就是特征值，对应于 λ_1 的一个特征向量是 $(1, 0)$，对应于 λ_2 的一个特征向量是 $(0, 1)$．从而方程的通解为

$$\boldsymbol{X}(t) = \alpha e^{\lambda_1 t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta e^{\lambda_2 t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

由于 $\lambda_1 < 0$，位于 x 轴上形如 $\alpha e^{\lambda_1 t}(1, 0)$ 的直线解在 $t \to \infty$ 时趋于 $(0, 0)$．这个坐标轴称为**稳定线**．由于 $\lambda_2 > 0$，位于 y 轴上解 $\beta e^{\lambda_2 t}(0, 1)$ 在 $t \to \infty$ 时，远离 $(0, 0)$．

这个坐标轴称为**不稳定线**. 由于在 t 增加时, $\boldsymbol{X}(t)$ 与 $(0, \beta e^{\lambda_2 t})$ 越来越近, 因而在 $t \to \infty$ 时, 其他的解 $(\alpha, \beta \neq 0)$ 都将沿不稳定线趋于 ∞. 而在负向时, 这些解都将沿稳定线趋于 ∞. ■

我们在图 3.1 中作出了该系统的**相图**. 所谓一个系统的相图就是指一个系统的一些有代表意义的解曲线在**相平面**$\mathbb{R}^2$ 上的图像. 系统的这种平衡点 (特征值满足 $\lambda_1 < 0 < \lambda_2$) 称为**鞍点**.

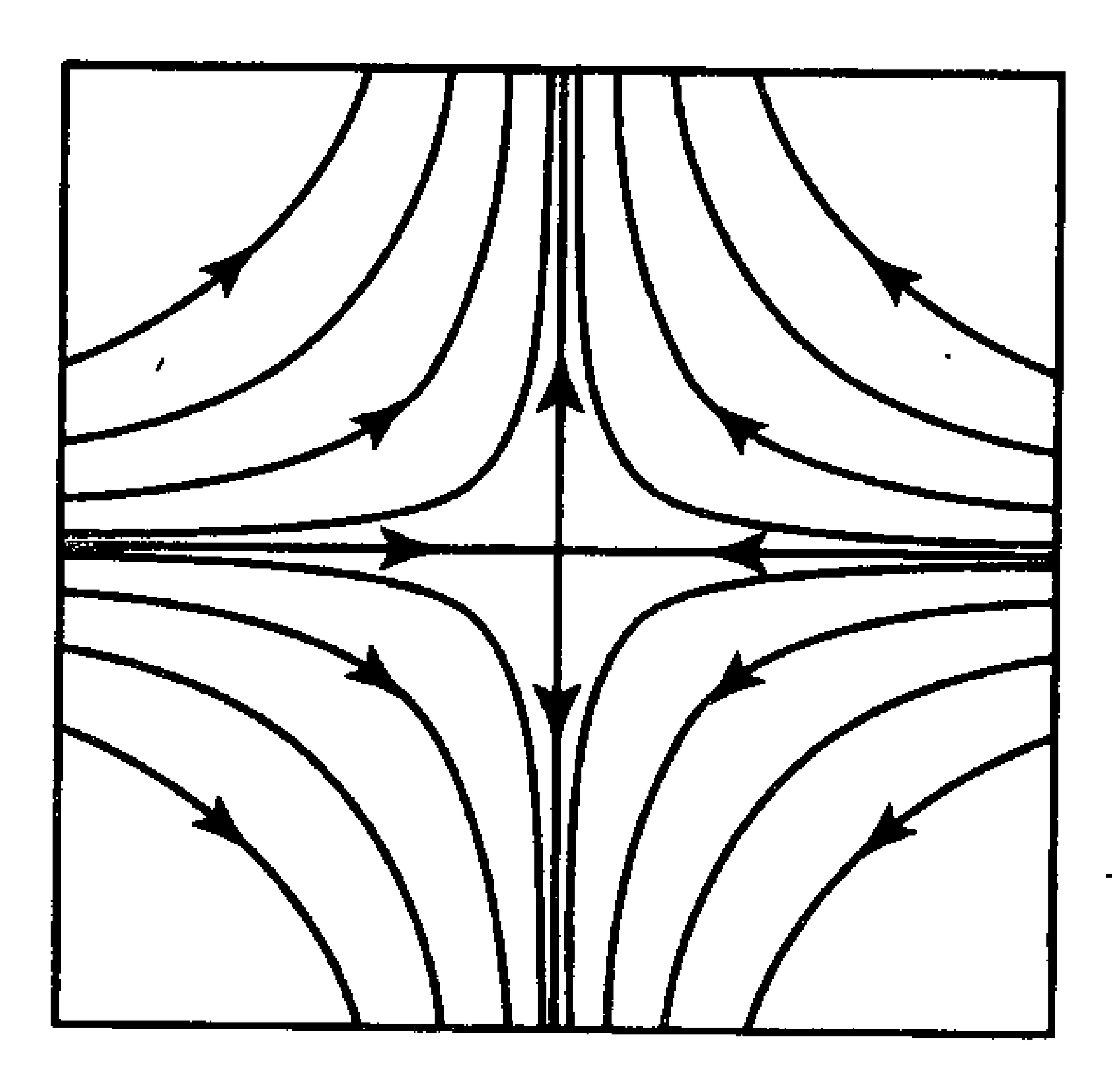


图 3.1　$x' = -x, y' = y$ 的鞍点相图

我们来看这种类型的一个稍微复杂一点的例子. 考虑系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{A} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}.$$

在第 2 章我们已经知道 $\boldsymbol{A}$ 的特征值为 ± 2. 对应于 $\lambda = 2$ 的特征向量是 $(3, 1)$, 而对应于 $\lambda = -2$ 的特征向量是 $(1, -1)$. 于是就有由形如

$$\boldsymbol{X}_1(t) = \alpha e^{2t} \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

的直线解构成的不稳定线, 当 $t \to \infty$ 时, 这些解都将远离原点. 形如

$$\boldsymbol{X}_2(t) = \beta e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

的直线解则构成了稳定线, 当 $t \to \infty$ 时, 这些解都将趋于原点. 根据线性叠加原理, 其他的解都具有形式

$$\boldsymbol{X}(t) = \alpha e^{2t} \begin{pmatrix} 3 \\ 1 \end{pmatrix} + \beta e^{-2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

(对某给定的 α,β). 注意, 如果 $\alpha\neq 0$, 则当 $t\to\infty$ 时, 我们有

$$\boldsymbol{X}(t)\sim\alpha e^{2t}\begin{pmatrix}3\\1\end{pmatrix}=\boldsymbol{X}_1(t),$$

而如果 $\beta\neq 0$, 则当 $t\to-\infty$ 时, 我们有

$$\boldsymbol{X}(t)\sim\beta e^{-2t}\begin{pmatrix}1\\-1\end{pmatrix}=\boldsymbol{X}_2(t).$$

于是当时间增加时, 系统的典型解都将接近 $\dot{\boldsymbol{X}}_1(t)$, 而当时间减少时, 它们将趋于 $\boldsymbol{X}_2(t)$(该系统的相图见图 3.2). 这正好与上一例子相似.

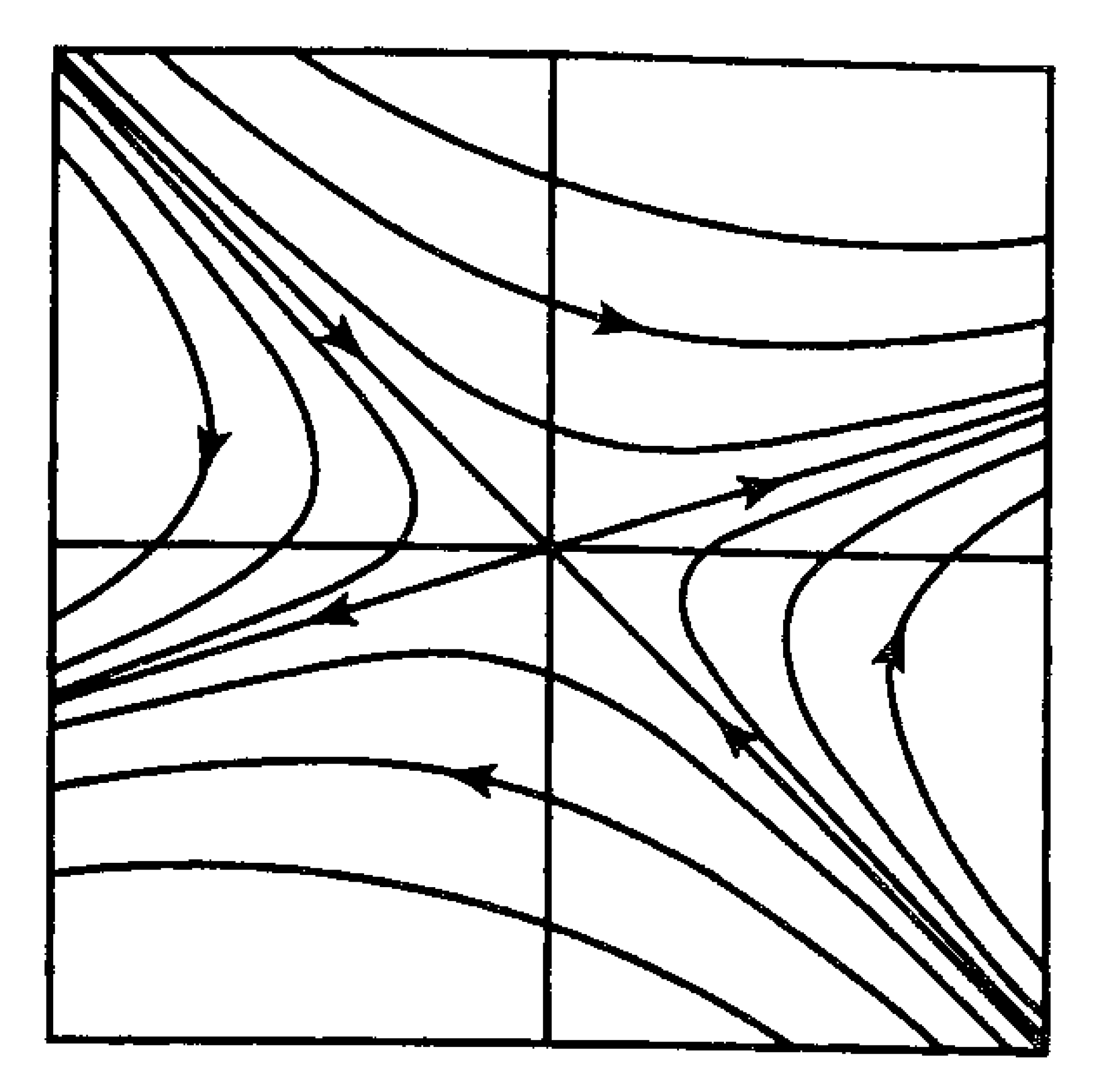


图 3.2 $x'=x+3y, y'=x-y$ 的鞍点相图

一般地, 当 $\boldsymbol{A}$ 具有一正一负的特征值时, 我们都可以找到类似的稳定线和不稳定线, 其上的解分别趋于或远离原点, 而其他的解在 $t\to\infty$ 时趋于不稳定线, 在 $t\to-\infty$ 时趋于稳定线.

例 (**汇点**) 现在考虑系统 $\boldsymbol{X}'=\boldsymbol{AX}$, 其中

$$\boldsymbol{A}=\begin{pmatrix}\lambda_1&0\\0&\lambda_2\end{pmatrix},$$

但是 $\lambda_1<\lambda_2<0$. 和前面一样, 我们可以找到两个直线解, 从而得到通解

$$\boldsymbol{X}(t)=\alpha e^{\lambda_1 t}\begin{pmatrix}1\\0\end{pmatrix}+\beta e^{\lambda_2 t}\begin{pmatrix}0\\1\end{pmatrix}.$$

与鞍点情形不同, 此时所有的解在 $t\to\infty$ 时都趋于 $(0,0)$. 现在要问: 它们以怎样的方式趋于原点? 为了回答这个问题, 我们来计算一个解的斜率 $\mathrm{d}y/\mathrm{d}x$ (假设

$\beta \neq 0$). 记

$$x(t) = \alpha e^{\lambda_1 t}$$
$$y(t) = \beta e^{\lambda_2 t}.$$

于是,

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{\lambda_2 \beta e^{\lambda_2 t}}{\lambda_1 \alpha e^{\lambda_1 t}} = \frac{\lambda_2 \beta}{\lambda_1 \alpha} e^{(\lambda_2 - \lambda_1)t}.$$

因为 $\lambda_2 - \lambda_1 > 0$, 从而这些斜率趋于 $\pm\infty$(假设 $\beta \neq 0$). 于是, 这些解将切于 y 轴趋于原点. ■

由于 $\lambda_1 < \lambda_2 < 0$, 我们称 λ_1 为强特征值, λ_2 为弱特征值. 此时之所以如此称呼是因为解的 x 坐标趋于 0 比其 y 坐标趋于 0 要快得多. 这就解释了为什么当解趋于原点时 (除了 λ_1 的特征向量所对应的直线上的解), 这些解会朝弱特征值所对应的解直线聚集.

图 3.3a 给出了该系统的相图. 此时平衡点称为**汇点**.

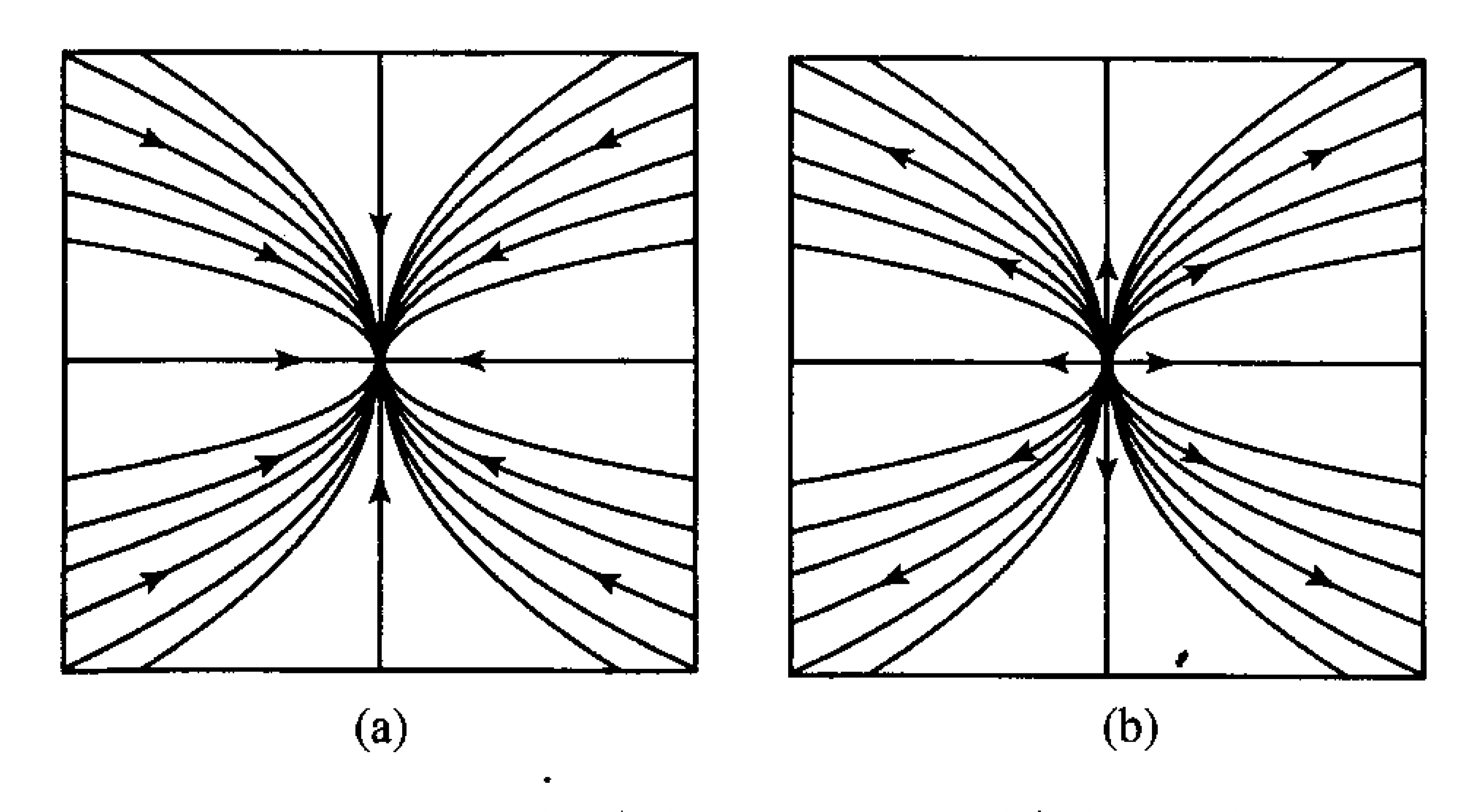


图 3.3 (a) 汇点的相图; (b) 源点的相图

一般地, 如果系统有特征值 $\lambda_1 < \lambda_2 < 0$, 对应的特征向量为 (u_1, u_2) 和 (v_1, v_2), 则它的通解为

$$\alpha e^{\lambda_1 t} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \beta e^{\lambda_2 t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

解的斜率为

$$\begin{aligned} \frac{dy}{dx} &= \frac{\lambda_1 \alpha e^{\lambda_1 t} u_2 + \lambda_2 \beta e^{\lambda_2 t} v_2}{\lambda_1 \alpha e^{\lambda_1 t} u_1 + \lambda_2 \beta e^{\lambda_2 t} v_1} \\ &= \left(\frac{\lambda_1 \alpha e^{\lambda_1 t} u_2 + \lambda_2 \beta e^{\lambda_2 t} v_2}{\lambda_1 \alpha e^{\lambda_1 t} u_1 + \lambda_2 \beta e^{\lambda_2 t} v_1} \right) \frac{e^{-\lambda_2 t}}{e^{-\lambda_2 t}} \\ &= \frac{\lambda_1 \alpha e^{(\lambda_1 - \lambda_2)t} u_2 + \lambda_2 \beta v_2}{\lambda_1 \alpha e^{(\lambda_1 - \lambda_2)t} u_1 + \lambda_2 \beta v_1}, \end{aligned}$$

它们将趋于 λ_2 的特征向量的斜率 v_2/v_1, 除非 $\beta=0$. 如果 $\beta=0$, 我们的解是特征值 λ_1 所对应的直线解. 因此, 所有的解 (除了那些强特征值所对应的直线解) 都与弱特征值所对应的直线解相切地趋于原点.

例 **(源点)**当矩阵

$$A=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$$

满足 $0<\lambda_2<\lambda_1$ 时, 对应的向量场可看成是上一个例子的负向量场. 其通解和相图都是一样的, 只是所有的解都沿着相同的路线远离 $(0,0)$(见图 3.3b). ■

现在, 可能有人会说我们所展示的例子过于简单了. 现在看来的确如此, 但是随后我们将看到, 任何具有不同实特征值的微分方程系统都可以通过坐标变换化成这种特殊形式.

最后, 当有一个特征值等于 0 时, 情况会有些特别. 我们已经知道, 此时有一条直线上的点全都是平衡点. 如果另一个特征值 λ 非零, 则 λ 的符号决定了其他的解是趋于这些平衡点还是远离这些平衡点 (见本章的习题 10 和习题 11).

3.2 复特征值

有时, 特征多项式的根会是复数, 与实情形类似, 我们称这些根为**复特征值**. 当矩阵 A 有复特征值时, 我们不再有直线解, 然而, 通过利用一些复数及复函数的技巧, 我们仍然可以像以前一样得到通解. 在下面的例子中, 我们将看到一般的过程是怎样的.

例 **(中心)**考虑系统 $X'=AX$, 其中

$$A=\begin{pmatrix}0 & \beta\\ -\beta & 0\end{pmatrix},$$

并且 $\beta\neq 0$. 其特征多项式为 $\lambda^2+\beta^2=0$, 于是特征值为虚数 $\pm\mathrm{i}\beta$. 如果不担心可能出现的复向量, 我们可以像以前一样去寻找与 $\lambda=\mathrm{i}\beta$ 相对应的特征向量. 这需要求解方程组

$$\begin{pmatrix}-\mathrm{i}\beta & \beta\\ -\beta & -\mathrm{i}\beta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}.$$

由于第二个方程是多余的, 上述方程组等价于 $\mathrm{i}\beta x=\beta y$. 于是得到一个复特征向量 $(1,\mathrm{i})$, 从而函数

$$X(t)=\mathrm{e}^{\mathrm{i}\beta t}\begin{pmatrix}1\\ \mathrm{i}\end{pmatrix}$$

为 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 相应的复解.

通常, 对一个实微分方程系统得到一个复解不是太合适, 但我们可以通过欧拉公式

$$\mathrm{e}^{\mathrm{i}\beta t} = \cos\beta t + \mathrm{i}\sin\beta t$$

来克服这一点. 利用欧拉公式, 可将解写成

$$\boldsymbol{X}(t) = \begin{pmatrix} \cos\beta t + \mathrm{i}\sin\beta t \\ \mathrm{i}(\cos\beta t + \mathrm{i}\sin\beta t) \end{pmatrix} = \begin{pmatrix} \cos\beta t + \mathrm{i}\sin\beta t \\ -\sin\beta t + \mathrm{i}\cos\beta t \end{pmatrix}.$$

将 $\boldsymbol{X}(t)$ 的实部和虚部分开, 可以写得更好些

$$\boldsymbol{X}(t) = \boldsymbol{X}_{\mathrm{Re}}(t) + \mathrm{i}\boldsymbol{X}_{\mathrm{Im}}(t),$$

其中

$$\boldsymbol{X}_{\mathrm{Re}}(t) = \begin{pmatrix} \cos\beta t \\ -\sin\beta t \end{pmatrix},\quad \boldsymbol{X}_{\mathrm{Im}}(t) = \begin{pmatrix} \sin\beta t \\ \cos\beta t \end{pmatrix}.$$

但是现在, 我们发现 $\boldsymbol{X}_{\mathrm{Re}}(t)$ 和 $\boldsymbol{X}_{\mathrm{Im}}(t)$ 都是原系统的(**实**)解. 为此, 我们来简单地验算一下:

$$\begin{aligned} \boldsymbol{X}'_{\mathrm{Re}}(t) + \mathrm{i}\boldsymbol{X}'_{\mathrm{Im}}(t) &= \boldsymbol{X}'(t) = \boldsymbol{A}\boldsymbol{X}(t) = \boldsymbol{A}(\boldsymbol{X}_{\mathrm{Re}}(t) + \mathrm{i}\boldsymbol{X}_{\mathrm{Im}}(t)) \\ &= \boldsymbol{A}\boldsymbol{X}_{\mathrm{Re}}(t) + \mathrm{i}\boldsymbol{A}\boldsymbol{X}_{\mathrm{Im}}(t). \end{aligned}$$

由于实部和虚部分别相等, 从而就有 $\boldsymbol{X}'_{\mathrm{Re}}(t) = \boldsymbol{A}\boldsymbol{X}_{\mathrm{Re}}(t)$ 和 $\boldsymbol{X}'_{\mathrm{Im}}(t) = \boldsymbol{A}\boldsymbol{X}_{\mathrm{Im}}(t)$, 这表明实部和虚部都是解. 进一步, 由于

$$\boldsymbol{X}_{\mathrm{Re}}(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix},\quad \boldsymbol{X}_{\mathrm{Im}}(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

它们的线性组合

$$\boldsymbol{X}(t) = c_1\boldsymbol{X}_{\mathrm{Re}}(t) + c_2\boldsymbol{X}_{\mathrm{Im}}(t)$$

就给出了任一初值问题的一个解, 其中 c_1 和 c_2 为任意常数.

我们断言上式也是方程的通解. 为此, 我们需要说明方程只有这些解. 否则, 假设

$$\boldsymbol{Y}(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$$

为另一个解, 考虑复函数 $f(t) = (u(t) + \mathrm{i}v(t))\mathrm{e}^{\mathrm{i}\beta t}$, 对它求导数并利用 $\boldsymbol{Y}(t)$ 是解的事实可得 $f'(t) = 0$. 于是, $u(t) + \mathrm{i}v(t)$ 是一个复常数乘上 $\mathrm{e}^{-\mathrm{i}\beta t}$, 由此可得 $\boldsymbol{Y}(t)$ 就是 $\boldsymbol{X}_{\mathrm{Re}}(t)$ 和 $\boldsymbol{X}_{\mathrm{Im}}(t)$ 的一个线性组合.

可以看到, 所有的这些解都是周期为 $2\pi/\beta$ 的周期函数, 事实上, 从系统的相图可以看出, 所有的解都在以原点为中心的圆周上. 当 $\beta > 0$ 时, 解沿圆周顺时针旋转, 而当 $\beta < 0$ 时则反时针旋转 (见图 3.4). 这种类型的系统称为一个**中心**. ■

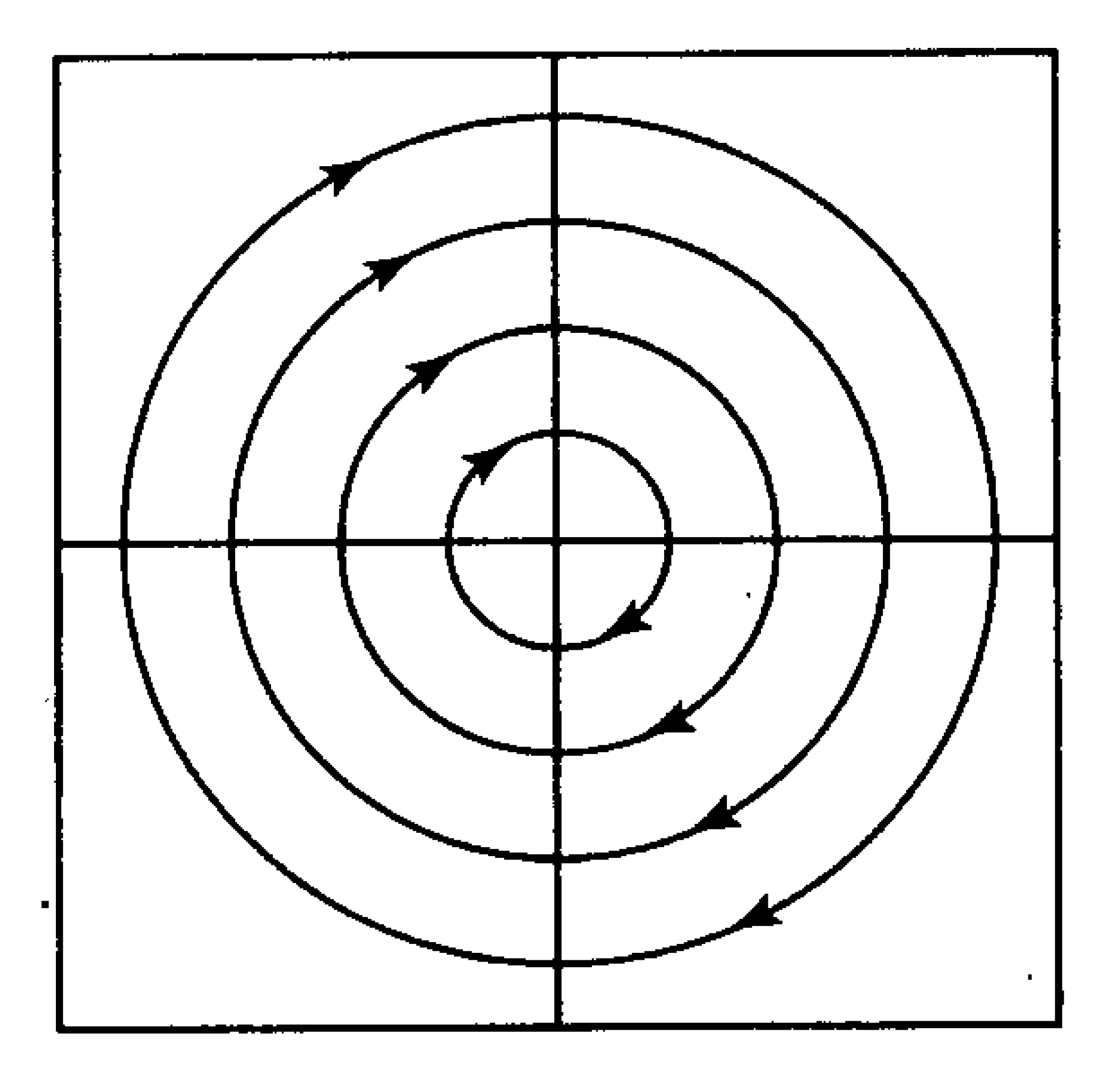


图 3.4 中心的相图

例 (**螺线汇点和螺线源点**)一般地, 考虑系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 其中

$$\boldsymbol{A} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix},$$

并且 $\alpha, \beta \neq 0$. 其特征多项式为 $\lambda^2 - 2\alpha\lambda + \alpha^2 + \beta^2$, 特征值为 $\lambda = \alpha \pm \mathrm{i}\beta$. 与 $\alpha + \mathrm{i}\beta$ 相对应的一个特征向量由方程

$$(\alpha - (\alpha + \mathrm{i}\beta))x + \beta y = 0$$

所确定. 从而 $(1, \mathrm{i})$ 仍然是一个特征向量, 由此可得如下的复解

$$\boldsymbol{X}(t) = \mathrm{e}^{(\alpha+\mathrm{i}\beta)t} \begin{pmatrix} 1 \\ \mathrm{i} \end{pmatrix} = \mathrm{e}^{\alpha t} \begin{pmatrix} \cos\beta t \\ -\sin\beta t \end{pmatrix} + \mathrm{i}\mathrm{e}^{\alpha t} \begin{pmatrix} \sin\beta t \\ \cos\beta t \end{pmatrix} = \boldsymbol{X}_{\mathrm{Re}}(t) + \mathrm{i}\boldsymbol{X}_{\mathrm{Im}}(t).$$

与刚才一样, $\boldsymbol{X}_{\mathrm{Re}}(t)$ 和 $\boldsymbol{X}_{\mathrm{Im}}(t)$ 都是系统的实解, 并且它们的初值条件是线性无关的. 这样我们就得到了通解

$$\boldsymbol{X}(t) = c_1\mathrm{e}^{\alpha t} \begin{pmatrix} \cos\beta t \\ -\sin\beta t \end{pmatrix} + c_2\mathrm{e}^{\alpha t} \begin{pmatrix} \sin\beta t \\ \cos\beta t \end{pmatrix}.$$

如果没有 $\mathrm{e}^{\alpha t}$ 这一项, 这些解将周期地缠绕在以原点为中心的圆周上, 而多了 $\mathrm{e}^{\alpha t}$ 这一项将使得解要么盘旋地进入原点 (当 $\alpha < 0$ 时), 要么盘旋地离开原点 (当 $\alpha > 0$ 时). 此时平衡点分别称为**螺线汇点**或**螺线源点**(见图 3.5). ■

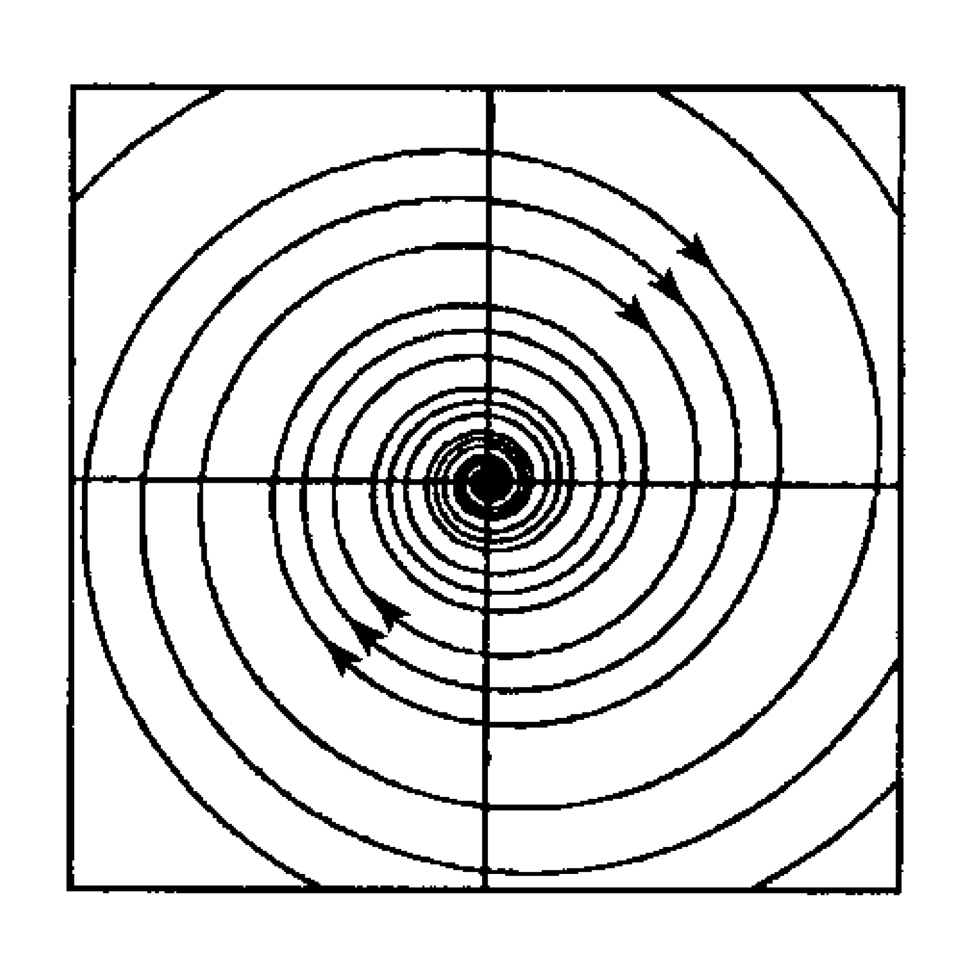
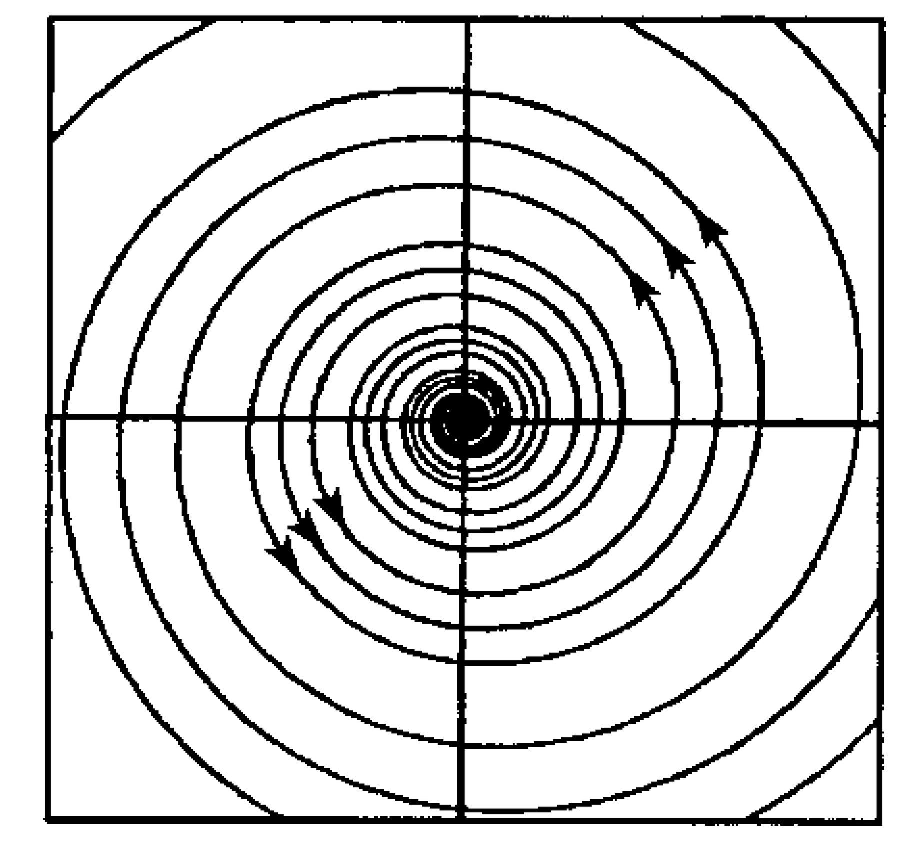



图 3.5 螺线汇点和螺线源点的相图

3.3 重特征值

现在剩下要讨论的情形就是 $\boldsymbol{A}$ 有重的实特征值情形. 它的一个简单形式就是 $\boldsymbol{A}$ 为对角矩阵

$$\boldsymbol{A} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

$\boldsymbol{A}$ 的两个特征值都是 λ. 此时, 对任给的 $\boldsymbol{V} \in \mathbb{R}^2$,

$$\boldsymbol{AV} = \lambda \boldsymbol{V},$$

因而任何非零向量都是特征向量. 于是任何解都可以写成

$$\boldsymbol{X}(t) = \alpha \mathrm{e}^{\lambda t} \boldsymbol{V}.$$

每一个解都在通过原点的直线上, 要么趋于原点 (当 $\lambda < 0$ 时), 要么远离原点 (当 $\lambda > 0$ 时). 因而, 这是一种容易的情形.

更有趣的情形是

$$\boldsymbol{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

的情形. 此时两个特征值仍然都等于 λ. 但此时只有一个线性无关的特征向量 $(1, 0)$. 从而其对应的直线解为

$$\boldsymbol{X}_1(t) = \alpha \mathrm{e}^{\lambda t} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

为了找到其他的解, 我们将系统写成

$$\begin{aligned} x' &= \lambda x + y \\ y' &= \lambda y. \end{aligned}$$

当 $y \neq 0$ 时, 则有

$$y(t) = \beta e^{\lambda t}.$$

于是关于 $x(t)$ 的微分方程就变成

$$x' = \lambda x + \beta e^{\lambda t}.$$

这是关于 $x(t)$ 的一个非自治一阶微分方程. 可能会有人猜测解的形式为 $e^{\lambda t}$, 但是其非自治项也是这种形式的. 可能你们在微积分课上已经知道, 最好假设解的可能形式为

$$x(t) = \alpha e^{\lambda t} + \mu t e^{\lambda t},$$

其中 α, μ 为常数. 这种技巧通常称为 "待定系数法". 将上式代入微分方程可得 $\mu = \beta$, 而 α 则是任意的. 从而系统的解可以写成

$$\alpha e^{\lambda t}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta e^{\lambda t}\begin{pmatrix} t \\ 1 \end{pmatrix}.$$

这事实上就是系统的通解 (见习题 12).

可以看出, 如果 $\lambda < 0$, 当 $t \to \infty$, 解的每一项都趋于 0. 对于 $\alpha e^{\lambda t}$ 和 $\beta e^{\lambda t}$ 两项, 这是显然的, 对于 $\beta t e^{\lambda t}$ 则可由罗必达法则立刻得到. 于是在 $t \to \infty$ 时, 所有的解都趋于 $(0,0)$. 而当 $\lambda > 0$ 时, 所有的解都远离 $(0,0)$, 见图 3.6. 事实上, 解都是沿特征向量 $(1,0)$ 的方向趋于或远离原点的 (见习题 7).

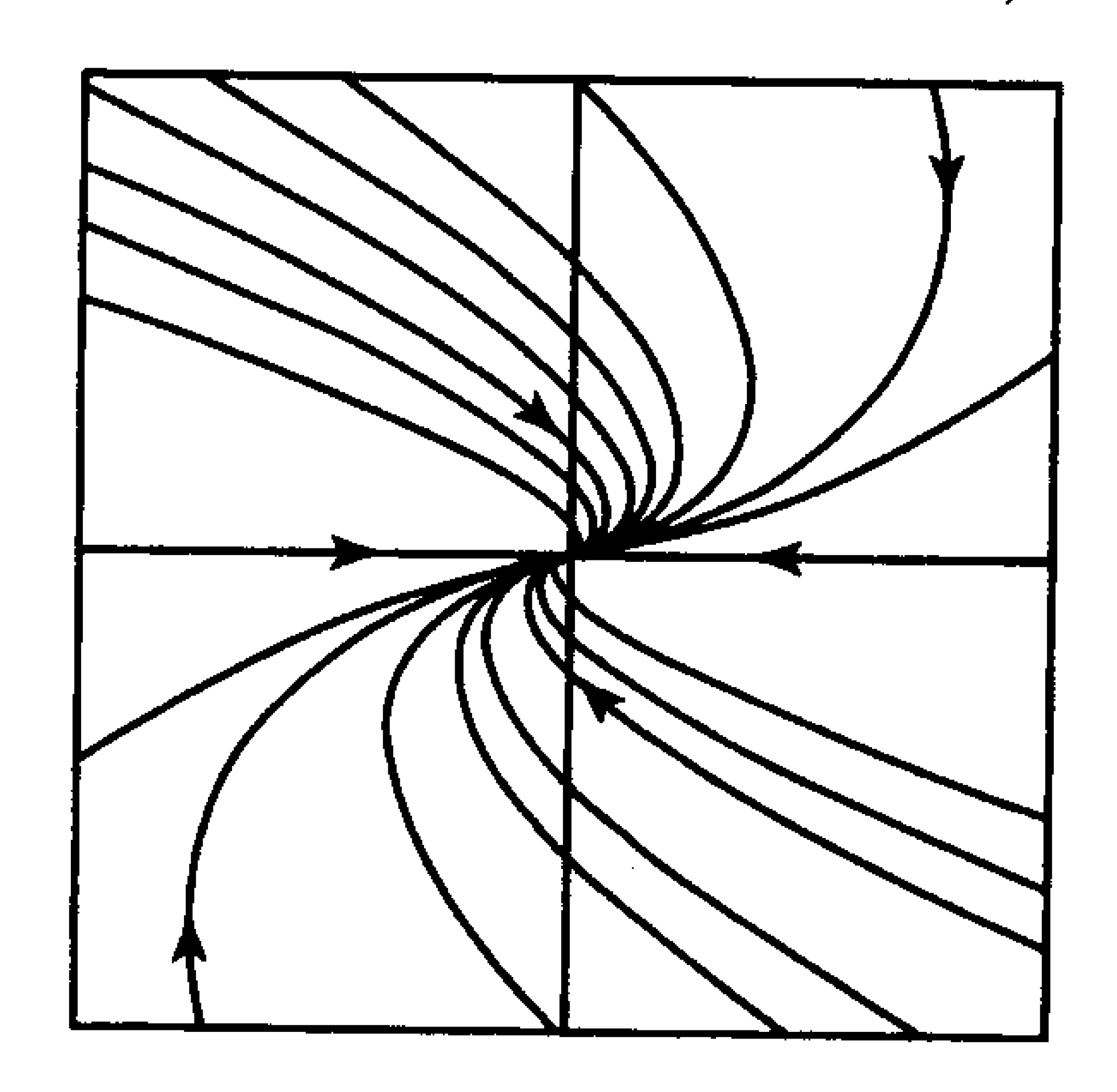


图 3.6 一个具有重负特征值系统的相图

3.4 坐标变换

在前面三节, 除去相图的不同外, 我们实际上只处理了以下三种类型的矩阵

$$\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}, \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

其中在第一种情形 λ 可能等于 μ.

任何这种形式的 2×2 矩阵称为**标准形**. 这种形式的系统似乎相当特别, 但事实并非如此. 任给线性系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 我们总可以通过 "坐标变换", 使得新系统的系数矩阵成为标准形, 从而变得容易求解. 下面我们就来做这件事.

$\mathbb{R}^2$ 上的一个**线性映射**(或**线性变换**) 是指一个如下形式的函数 $T:\mathbb{R}^2\to\mathbb{R}^2$:

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}.$$

也就是说, T 的作用就是用 2×2 矩阵

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

去乘任一向量. 因而我们认为线性映射和它对应的矩阵是可以互换使用的, 从而也写成

$$\boldsymbol{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

我们希望这一点点的不精确不会带来混淆.

现在假设 $\boldsymbol{T}$ 是**可逆**的, 也就是说矩阵 $\boldsymbol{T}$ 有**逆矩阵** $\boldsymbol{S}$ 满足 $\boldsymbol{T}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{T} = \boldsymbol{I}$, 其中 $\boldsymbol{I}$ 为 2×2 单位矩阵. 习惯上我们将矩阵 $\boldsymbol{T}$ 的逆记为 $\boldsymbol{T}^{-1}$. 容易验证当 $\det\boldsymbol{T}\neq 0$ 时, 矩阵

$$\boldsymbol{S} = \frac{1}{\det\boldsymbol{T}}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

就是 $\boldsymbol{T}^{-1}$. 如果 $\det\boldsymbol{T} = 0$, 由第 2 章可知, 有无穷多个向量 (x,y) 使得

$$\boldsymbol{T}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

于是在这种情形没有逆矩阵, 否则的话, 对所有满足上式的向量, 我们都有

$$\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{T}^{-1}\boldsymbol{T}\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{T}^{-1}\begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

这样我们就证明了:

命题 2×2 矩阵 $\boldsymbol{T}$ 为可逆的当且仅当 $\det\boldsymbol{T}\neq 0$. ■

现在我们来考虑系统

$$\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y},$$

(而不是线性系统 $\boldsymbol{X}'=\boldsymbol{AX}$), 其中 $\boldsymbol{T}$ 是某一可逆线性映射. 可见, 如果 $\boldsymbol{Y}(t)$ 是新系统的一个解, 则 $\boldsymbol{X}(t)=\boldsymbol{TY}(t)$ 就是 $\boldsymbol{X}'=\boldsymbol{AX}$ 的一个解, 理由如下:

$$(\boldsymbol{TY}(t))'=\boldsymbol{TY}'(t)=\boldsymbol{T}(\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{Y}(t)=\boldsymbol{A}(\boldsymbol{TY}(t)).$$

也就是说, 线性映射 $\boldsymbol{T}$ 将 $\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{Y}$ 的解变换成了 $\boldsymbol{X}'=\boldsymbol{AX}$ 的解. 反过来, $\boldsymbol{T}^{-1}$ 则将 $\boldsymbol{X}'=\boldsymbol{AX}$ 的解变成了 $\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{Y}$ 的解.

从而 $\boldsymbol{T}$ 可以看成是一个坐标变换, 它将一个给定的线性系统变成另外一个系数矩阵不同的线性系统. 我们希望的是, 对一给定系统, 找到一个线性映射 $\boldsymbol{T}$, 使得经过变换得到的系统 $\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{Y}$ 容易求解. 你们也许可以猜到, 我们总可以找到一个线性映射将一个给定的线性系统变成标准形中的一个.

例 (**实特征值**) 假设矩阵 $\boldsymbol{A}$ 有两个不同的实特征值 λ_1 和 λ_2, 相应的特征向量分别为 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$. 记 $\boldsymbol{T}$ 为这样的矩阵, 它的列向量是 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$. 于是就有 $\boldsymbol{TE}_j=\boldsymbol{V}_j, j=1,2$, 其中 $\boldsymbol{E}_1$ 和 $\boldsymbol{E}_2$ 构成 $\mathbb{R}^2$ 的标准基. 这也就是 $\boldsymbol{T}^{-1}\boldsymbol{V}_j=\boldsymbol{E}_j$, 从而我们有

$$(\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{E}_j=\boldsymbol{T}^{-1}\boldsymbol{AV}_j=\boldsymbol{T}^{-1}(\lambda_j\boldsymbol{V}_j)=\lambda_j\boldsymbol{T}^{-1}\boldsymbol{V}_j=\lambda_j\boldsymbol{E}_j.$$

因而矩阵 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 为标准形

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix},$$

而相应的系统容易求解. ■

例 我们来看一个更具体的例子, 设

$$\boldsymbol{A}=\begin{pmatrix}-1 & 0\\ 1 & -2\end{pmatrix}.$$

其特征方程为 $\lambda^2+3\lambda+2=0$, 由此可得特征值为 $\lambda=-1$ 和 $\lambda=-2$. 通过解方程

$$(\boldsymbol{A}+\boldsymbol{I})\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0 & 0\\ 1 & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix},$$

可以求得与 $\lambda=-1$ 相对应的一个特征向量为 $(1,1)$; 类似地, 与 $\lambda=-2$ 相对应的一个特征向量为 $(0,1)$.

于是我们就得到了一对直线解, 在 $t\to\infty$ 时, 它们都趋于原点. 对应于弱特征值的直线解在直线 $y=x$ 上, 而对应于强特征值的直线解则在 y 轴上, 所有其他的解都以切于直线 $y=x$ 的方式趋于原点.

为了将这个系统化为标准形, 取 $\boldsymbol{T}$ 为以特征向量为列向量的矩阵:

$$\boldsymbol{T}=\begin{pmatrix}1&0\\1&1\end{pmatrix},$$

从而

$$\boldsymbol{T}^{-1}=\begin{pmatrix}1&0\\-1&1\end{pmatrix}.$$

最后可以算得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}-1&0\\0&-2\end{pmatrix},$$

故 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 为标准形. 系统 $\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的通解为

$$\boldsymbol{Y}(t)=\alpha\mathrm{e}^{-t}\begin{pmatrix}1\\0\end{pmatrix}+\beta\mathrm{e}^{-2t}\begin{pmatrix}0\\1\end{pmatrix},$$

从而 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的通解为

$$\begin{aligned}\boldsymbol{T}\boldsymbol{Y}(t)&=\begin{pmatrix}1&0\\1&1\end{pmatrix}\left(\alpha\mathrm{e}^{-t}\begin{pmatrix}1\\0\end{pmatrix}+\beta\mathrm{e}^{-2t}\begin{pmatrix}0\\1\end{pmatrix}\right)\\&=\alpha\mathrm{e}^{-t}\begin{pmatrix}1\\1\end{pmatrix}+\beta\mathrm{e}^{-2t}\begin{pmatrix}0\\1\end{pmatrix}.\end{aligned}$$

于是, 线性映射 $\boldsymbol{T}$ 将系统

$$\boldsymbol{Y}'=\begin{pmatrix}-1&0\\0&-2\end{pmatrix}\boldsymbol{Y}$$

的相图映成了 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的相图 (见图 3.7). ■

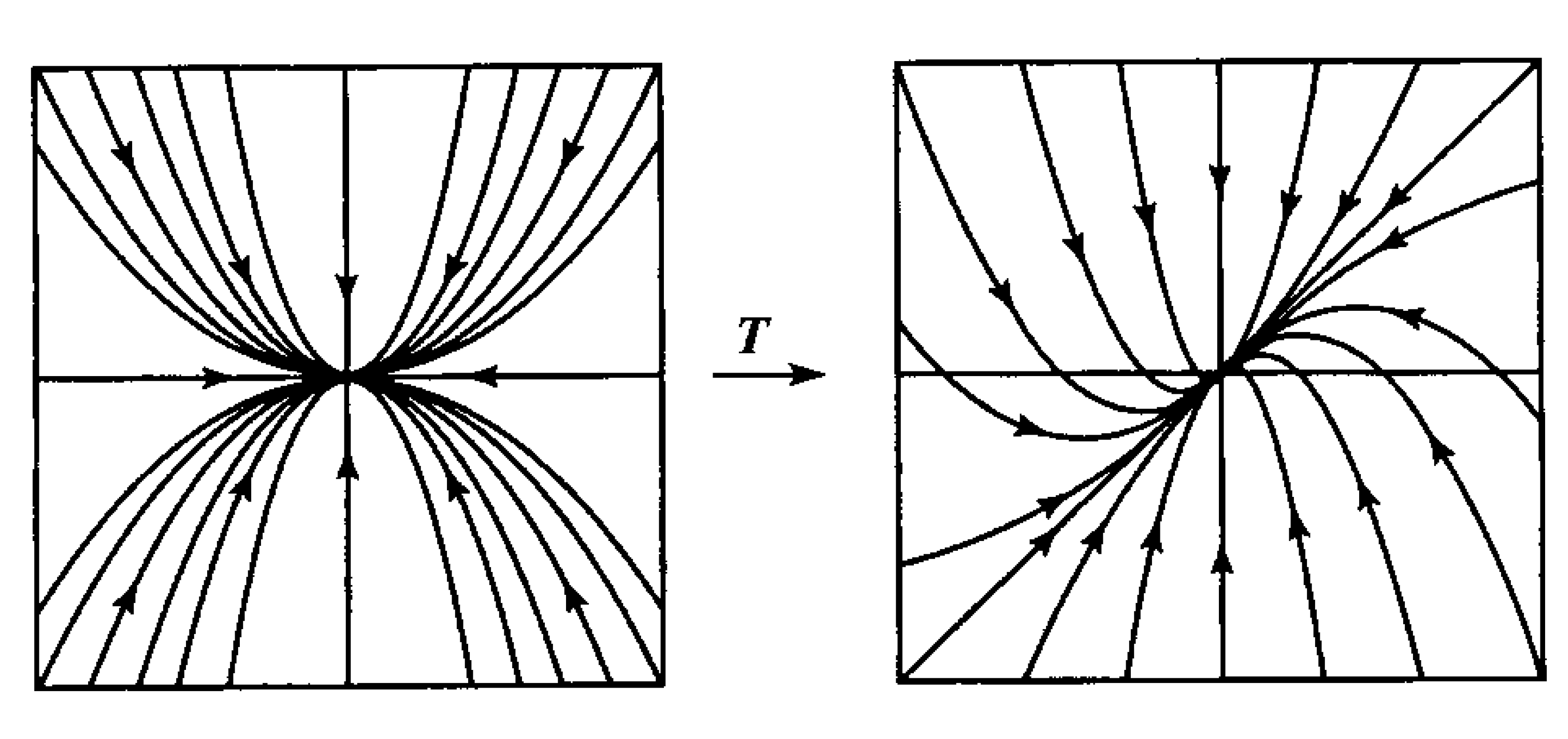


图 3.7　(实) 汇点情形的变量替换 $\boldsymbol{T}$

一点说明：一旦有了特征值和特征向量，我们可以简单地写出通解，也就是说此时我们并不真的需要将一个具体的系统化为标准形，而之所以做这额外的一步，是因为当我们试图将所有的线性系统分类时，系统的标准形将会极大地简化这一过程.

例 (**复特征值**) 现在假设 $\boldsymbol{A}$ 有复特征值 $\alpha \pm i\beta$，其中 $\beta \neq 0$. 于是，对应于 $\alpha + i\beta$，我们可以找到复特征向量 $\boldsymbol{V}_1 + i\boldsymbol{V}_2$，其中 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 都是实向量. 我们断言 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 是 $\mathbb{R}^2$ 中线性无关的向量. 否则，对某一 $c \in \mathbb{R}$，有 $\boldsymbol{V}_1 = c\boldsymbol{V}_2$. 于是就有

$$\boldsymbol{A}(\boldsymbol{V}_1 + i\boldsymbol{V}_2) = (\alpha + i\beta)(\boldsymbol{V}_1 + i\boldsymbol{V}_2) = (\alpha + i\beta)(c + i)\boldsymbol{V}_2.$$

另一方面，我们又有

$$\boldsymbol{A}(\boldsymbol{V}_1 + i\boldsymbol{V}_2) = (c + i)\boldsymbol{A}\boldsymbol{V}_2.$$

于是，我们可以推得 $\boldsymbol{A}\boldsymbol{V}_2 = (\alpha + i\beta)\boldsymbol{V}_2$. 由于左边是实向量而右边是复向量，这不可能.

由于 $\boldsymbol{V}_1 + i\boldsymbol{V}_2$ 是属于 $\alpha + i\beta$ 的特征向量，我们有

$$\boldsymbol{A}(\boldsymbol{V}_1 + i\boldsymbol{V}_2) = (\alpha + i\beta)(\boldsymbol{V}_1 + i\boldsymbol{V}_2).$$

由这个向量方程的实部和虚部分别相等可得，

$$\begin{aligned}\boldsymbol{A}\boldsymbol{V}_1 &= \alpha\boldsymbol{V}_1 - \beta\boldsymbol{V}_2\\ \boldsymbol{A}\boldsymbol{V}_2 &= \beta\boldsymbol{V}_1 + \alpha\boldsymbol{V}_2.\end{aligned}$$

设 $\boldsymbol{T}$ 是以 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_2$ 为列向量的矩阵，于是就有 $\boldsymbol{T}\boldsymbol{E}_j = \boldsymbol{V}_j, j = 1, 2$. 对于矩阵 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$，我们有

$$(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{E}_1 = \boldsymbol{T}^{-1}(\alpha\boldsymbol{V}_1 - \beta\boldsymbol{V}_2) = \alpha\boldsymbol{E}_1 - \beta\boldsymbol{E}_2,$$

同样还有

$$(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{E}_2 = \beta\boldsymbol{E}_1 + \alpha\boldsymbol{E}_2.$$

从而矩阵 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 为标准形

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

我们已经看到当 $\alpha < 0, \alpha = 0$ 和 $\alpha > 0$ 时，系统 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的相图分别对应于螺线汇点、中心和螺线源点. 从而在经过坐标变换 $\boldsymbol{T}$ 之后，$\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的相图也与其中之一等价. ■

例　(另一个调和振子)考虑二阶方程

$$x'' + 4x = 0,$$

这对应于一个质量为 1, 弹性系数为 4 的无阻尼调和振子. 作为一个系统, 我们有

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \boldsymbol{X} = \boldsymbol{AX},$$

其特征方程为

$$\lambda^2 + 4 = 0,$$

从而特征值为 $\pm 2\mathrm{i}$. 与 $\lambda = 2\mathrm{i}$ 相对应的一个特征向量是系统

$$\begin{aligned} -2\mathrm{i}x + y &= 0 \\ -4x - 2\mathrm{i}y &= 0 \end{aligned}$$

的一个解. 可取 $(1, 2\mathrm{i})$ 为这样的一个解, 从而可以得到一个复解

$$\mathrm{e}^{2\mathrm{i}t} \begin{pmatrix} 1 \\ 2\mathrm{i} \end{pmatrix}.$$

将这个解分成实部和虚部后, 可以得到通解为

$$\boldsymbol{X}(t) = c_1 \begin{pmatrix} \cos 2t \\ -2\sin 2t \end{pmatrix} + c_2 \begin{pmatrix} \sin 2t \\ 2\cos 2t \end{pmatrix}.$$

于是振子的位置由

$$x(t) = c_1 \cos 2t + c_2 \sin 2t$$

给出, 这是一个周期为 π 的周期函数.

现在, 设 $\boldsymbol{T}$ 为这样的矩阵, 它的两个列向量分别为特征向量 $(1, 2\mathrm{i})$ 的实部和虚部, 即

$$\boldsymbol{T} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

从而很容易地算得

$$\boldsymbol{T}^{-1}\boldsymbol{AT} = \begin{pmatrix} 0 & 2 \\ -2 & 0 \end{pmatrix},$$

这是一个标准形. 这两个系统的相图见图 3.8. 可以看到, $\boldsymbol{T}$ 将系统 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{AT})\boldsymbol{Y}$ 的圆形解变成了 $\boldsymbol{X}' = \boldsymbol{AX}$ 的椭圆形解. ■

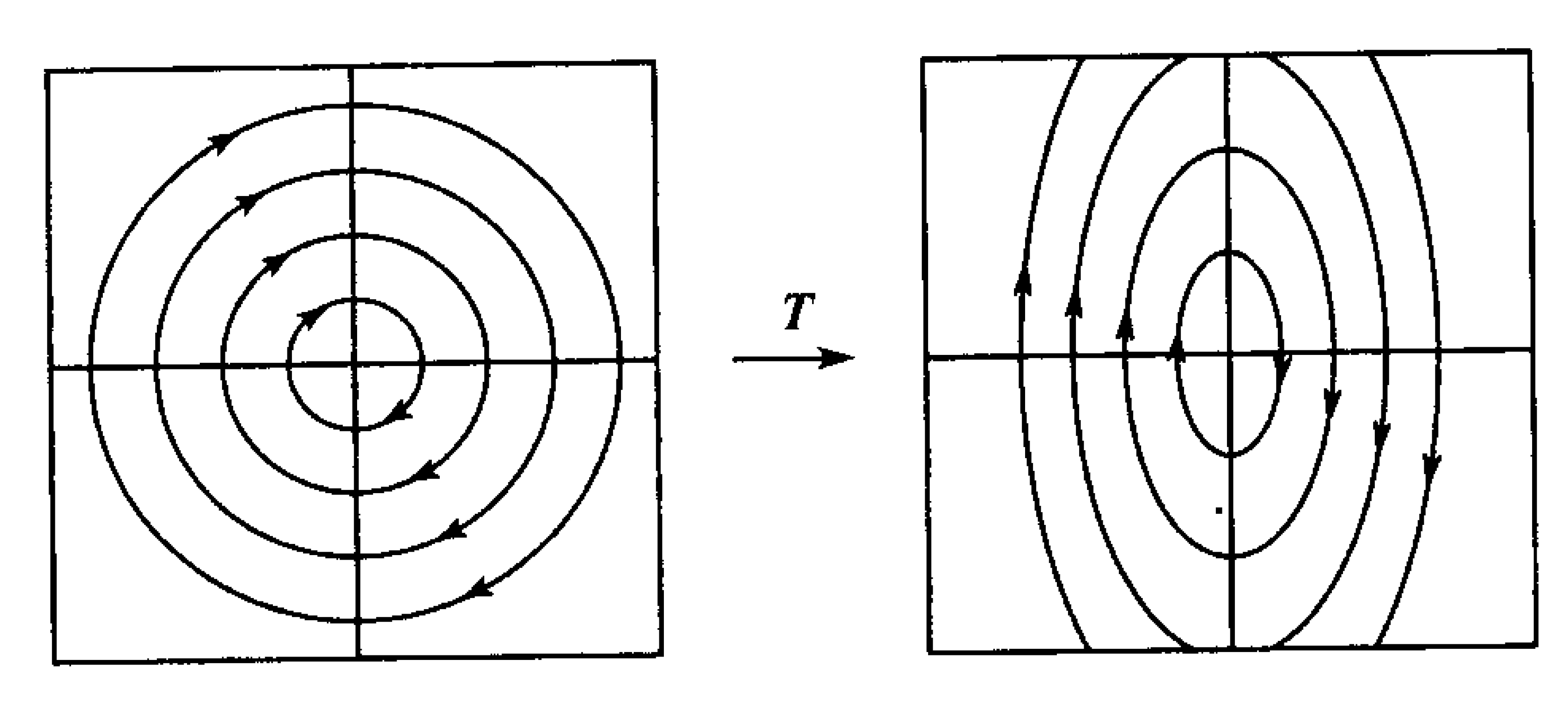


图 3.8　中心情形的变量替换 T

例　(**重特征值**)假设 A 只有一个特征值 λ. 如果 A 有一对线性无关的特征向量, 则它的形式必为

$$A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix},$$

此时系统 $X' = AX$ 容易求解 (见习题 15).

对于复杂些的情形, 假设 V 是一个特征向量, 其他的特征向量都是 V 的倍数, 任取向量 W 使得 V 和 W 是线性无关的, 于是存在常数 $\mu, \nu \in \mathbb{R}$, 使得

$$AW = \mu V + \nu W,$$

注意此时 $\mu \neq 0$, 因为否则的话, 我们就有了第二个线性无关的特征向量 W(对应于特征值 ν). 我们断言 $\nu = \lambda$. 如果 $\nu - \lambda \neq 0$, 经过计算可得

$$A\left(W + \left(\frac{\mu}{\nu - \lambda}\right)V\right) = \nu\left(W + \left(\frac{\mu}{\nu - \lambda}\right)V\right),$$

这说明 ν 是与 λ 不同的第二个特征值. 从而我们必有 $\nu = \lambda$.

最后, 令 $U = (1/\mu)W$, 则

$$AU = V + \frac{\lambda}{\mu}W = V + \lambda U.$$

因而如果定义 T 使得 $TE_1 = V, TE_2 = U$, 则可得

$$T^{-1}AT = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

这正是我们期望得到的. 从而, $X' = AX$ 经过坐标变换后也化成了标准形. ■

习　题

1. 在图 3.9 中有 6 个相图, 请将这些相图与下面的线性系统一一对应起来:

(a) $\begin{pmatrix} 3 & 5 \\ -2 & -2 \end{pmatrix}$ (b) $\begin{pmatrix} -3 & -2 \\ 5 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 3 & -2 \\ 5 & -2 \end{pmatrix}$

(d) $\begin{pmatrix} -3 & 5 \\ -2 & 3 \end{pmatrix}$ (e) $\begin{pmatrix} 3 & 5 \\ -2 & -3 \end{pmatrix}$ (f) $\begin{pmatrix} -3 & 5 \\ -2 & 2 \end{pmatrix}$

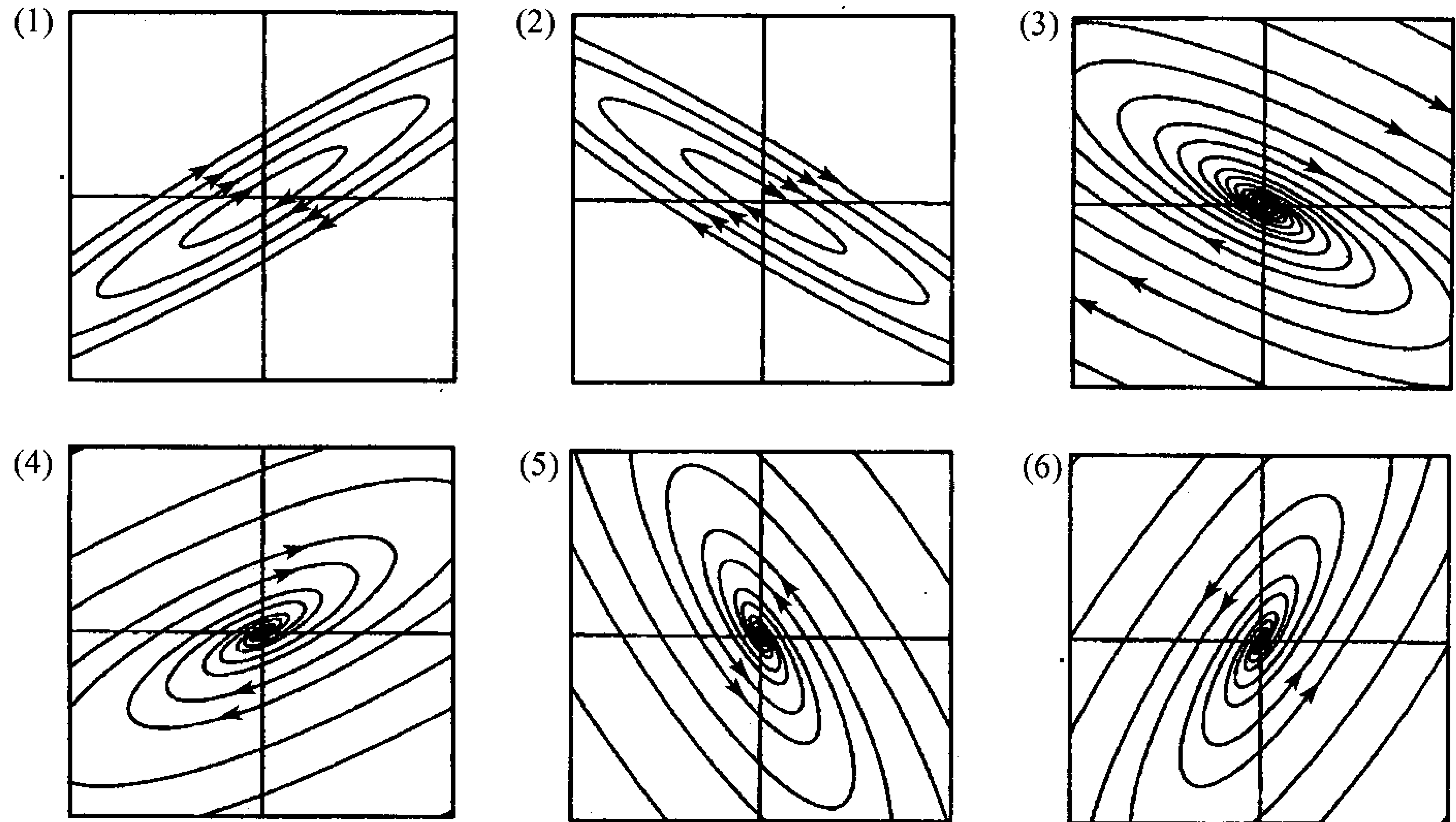


图 3.9 将这些相图与习题 1 中的系统对应

2. 对下面的每个系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$,
   (a) 找出 $\boldsymbol{A}$ 的特征值和特征向量.
   (b) 找出矩阵 $\boldsymbol{T}$ 把 $\boldsymbol{A}$ 化为标准形.
   (c) 找出 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 和 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的通解.
   (d) 简略地作出两个系统的相图.

(i) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (ii) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ (iii) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}$

(iv) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$ (v) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ -1 & -3 \end{pmatrix}$ (vi) $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

3. 找出下面调和振子方程的通解:
   (a) $x'' + x' + x = 0$ (b) $x'' + 2x' + x = 0$
4. 考虑调和振子系统

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 \\ -k & -b \end{pmatrix} \boldsymbol{X},$$

其中 $b \geqslant 0, k > 0$, 并且质量 $m = 1$.

(a) 当 k, b 取什么值时, 系统具有复特征值? 重特征值? 不同的实特征值?

(b) 在每种情形下, 找出系统的通解.

(c) 在每种情形下, 当质点在初始位置 $x = 1$ 以零速度释放时, 描述质点的运动.

5. 简略地作出系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的相图, 其中

$$\boldsymbol{A} = \begin{pmatrix} a & 1 \\ 2a & 2 \end{pmatrix}.$$

当 a 取什么值时会出现分岔? 描述在分岔点上下的 a 所对应系统的相图.

6. 考虑系统

$$\boldsymbol{X}' = \begin{pmatrix} 2a & b \\ b & 0 \end{pmatrix} \boldsymbol{X}.$$

在 ab 平面上, 描绘出不同标准形所对应的区域.

7. 考虑系统

$$\boldsymbol{X}' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \boldsymbol{X},$$

其中 $\lambda \neq 0$. 证明当 $\lambda < 0$ $(\lambda > 0)$ 时, 所有的解都以切于向量 $(1,0)$ 的方式趋于 (远离) 原点.

8. 找出所有具有纯虚特征值的 2×2 矩阵, 即, 找出一些条件, 使得当矩阵的元素满足这些条件时, 矩阵具有纯虚特征值.

9. 找出一个可计算的条件, 使得当条件满足时, 如果矩阵 $\boldsymbol{A}$ 有虚部非零的复特征值, 则 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的解绕原点反时针旋转.

10. 考虑系统

$$\boldsymbol{X}' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \boldsymbol{X},$$

其中 $a + d \neq 0$ 但 $ad - bc = 0$. 找出这个系统的通解并简略地画出相图.

11. 找出系统

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \boldsymbol{X}$$

的通解并完整地描述其相图.

12. 证明

$$\alpha e^{\lambda t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \beta e^{\lambda t} \begin{pmatrix} t \\ 1 \end{pmatrix}$$

是系统

$$X' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} X$$

的通解.

13. 证明任一 2×2 矩阵 $\boldsymbol{A}$ 总满足它自己的特征方程, 即, 如果 $\boldsymbol{A}$ 的特征方程为 $\lambda^2 + \alpha\lambda + \beta = 0$, 则矩阵 $\boldsymbol{A}^2 + \alpha \boldsymbol{A} + \beta \boldsymbol{I}$ 为 0 矩阵.
14. 假设 2×2 矩阵 $\boldsymbol{A}$ 有重特征值 λ. 利用习题 13 的结论, 证明: 任给 $\boldsymbol{V} \in \mathbb{R}^2$, 要么 $\boldsymbol{V}$ 是 $\boldsymbol{A}$ 的一个特征向量, 要么 $(\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{V}$ 为 $\boldsymbol{A}$ 的一个特征向量.
15. 假设矩阵 $\boldsymbol{A}$ 具有重特征值 λ, 并且有一对线性无关的特征向量. 证明

$$\boldsymbol{A} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}.$$

16. 考虑 (非线性) 系统

$$x' = |y|$$
$$y' = -x.$$

利用本章的方法描述系统的相图.

第 4 章 平面系统的分类

在本章, 我们将用一种动力系统的观点对迄今所得到的结论进行总结. 这意味着我们将至少得到 2×2 自治线性系统所有可能行为的一本完整字典. 我们先通过迹-行列式平面以几何方式给出一种字典, 另一种字典则要更具有动力系统味道, 此时需要用到共轭系统的概念.

4.1 迹-行列式平面

对一个矩阵

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

我们知道它的特征值就是它的特征方程

$$\lambda^2 - (a+d)\lambda + (ad - bc) = 0$$

的根. 特征方程的常数项就是 $\det \boldsymbol{A}$, λ 的系数也有一个名字 ——$a + d$ 称为 $\boldsymbol{A}$ 的**迹**, 记为 $\operatorname{tr} \boldsymbol{A}$.

于是特征值满足方程

$$\lambda^2 - (\operatorname{tr} \boldsymbol{A})\lambda + \det \boldsymbol{A} = 0,$$

由此可以解得

$$\lambda_{\pm} = \frac{1}{2}\left(\operatorname{tr} \boldsymbol{A} \pm \sqrt{(\operatorname{tr} \boldsymbol{A})^2 - 4 \det \boldsymbol{A}}\right).$$

注意到 $\lambda_+ + \lambda_- = \operatorname{tr} \boldsymbol{A}$, $\lambda_+ \lambda_- = \det \boldsymbol{A}$, 于是 $\boldsymbol{A}$ 的迹就是它的两个特征值的和, 而行列式则是两个特征值的乘积. 我们还将记 $T = \operatorname{tr} \boldsymbol{A}$, $D = \det \boldsymbol{A}$. 知道了 T 和 D 也就知道了 $\boldsymbol{A}$ 的两个特征值, 从而也就知道了 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 解的几乎所有几何特征. 例如, 从 T 和 D 的取值, 我们可以知道解是盘旋地趋于还是远离原点, 是否存在一个中心, 等等.

我们将通过在**迹-行列式平面**上画一个图来形象地给出这种分类. 在这个图中, 迹为 T, 行列式为 D 的矩阵对应于一个坐标为 (T, D) 的点. 该点在 TD 平面上的位置就确定了相图的几何特征. 例如, $T^2 - 4D$ 的符号可以告诉我们:

(1) 当 $T^2 - 4D < 0$ 时, 特征值是具有非零虚部的复数;

(2) 当 $T^2-4D>0$ 时, 是不同的实特征值;

(3) 当 $T^2-4D=0$ 时, 是实的重特征值.

从而, 从 TD 平面上 (T,D) 点相对于抛物线 $T^2-4D=0$ 的位置就可以得到 $\boldsymbol{A}$ 的特征值的代数信息.

从相平面去看, 我们可以说得更多些. 如果 $T^2-4D<0$, 则特征值的实部为 $T/2$, 于是,

(1) 当 $T<0$ 时, 为螺线汇点; (2) 当 $T>0$ 时, 为螺线源点; (3) 当 $T=0$ 时, 为中心.

如果 $T^2-4D>0$, 同样可进行进一步的分析. 在这一区域, 两个特征值都是实的. 如果 $D<0$, 则对应于鞍点, 由于 D 是两个特征值的乘积, 从而一个为正, 而另一个为负. 我们也可以用下面的计算来说明. 假设 $D<0$, 则有

$$T^2<T^2-4D,$$

从而

$$\pm T<\sqrt{T^2-4D}.$$

由此可得

$$T+\sqrt{T^2-4D}>0,\quad T-\sqrt{T^2-4D}<0$$

从而两个特征值都是实的, 并且符号不同. 当 $D>0$ 且 $T<0$ 时, 则两个特征值

$$T\pm\sqrt{T^2-4D}<0,$$

因而这对应于 (实的) 汇点. 类似地, $D>0, T>0$ 则对应于 (实的) 源点.

当 $D=0, T\neq 0$ 时, 我们得到一个零特征值, 而当 $D=T=0$ 时, 两个特征值都为零.

将所有这些信息一点点在 TD 平面上描绘出来, 我们就得到了所有不同线性系统的一个生动的概观. 上面的那些方程将 TD 平面分成了一些不同的区域, 每一个区域对应于一类特定的系统, 见图 4.1. 这就得到了所有 2×2 线性系统的一个几何分类.

下面我们要依次给些注记. 首先, 迹-行列式平面只是一个实际上的四维空间的二维表示: 2×2 矩阵是由四个参数, 即矩阵的四个元素所确定的. 因而 TD 平面上的每一个点都对应于无穷多个不同的矩阵. 虽然每一个矩阵都有相同的特征值, 但相图却可能会有些许不同, 例如, 对于中心、螺线汇点和螺线源点, 旋转的方向可以不同, 对于重特征值情形, 线性无关的特征向量的个数可以是 1 或 2.

我们也可以将迹-行列式平面看成平面线性系统的某种分岔图. 一个线性系统的单参数族对应于 TD 平面上的一条曲线, 当这条曲线穿过 T 轴、D 轴的正半

轴或者抛物线 $T^2-4D=0$ 时, 线性系统的相图就会产生分岔: 相图的几何形状将有大的变化.

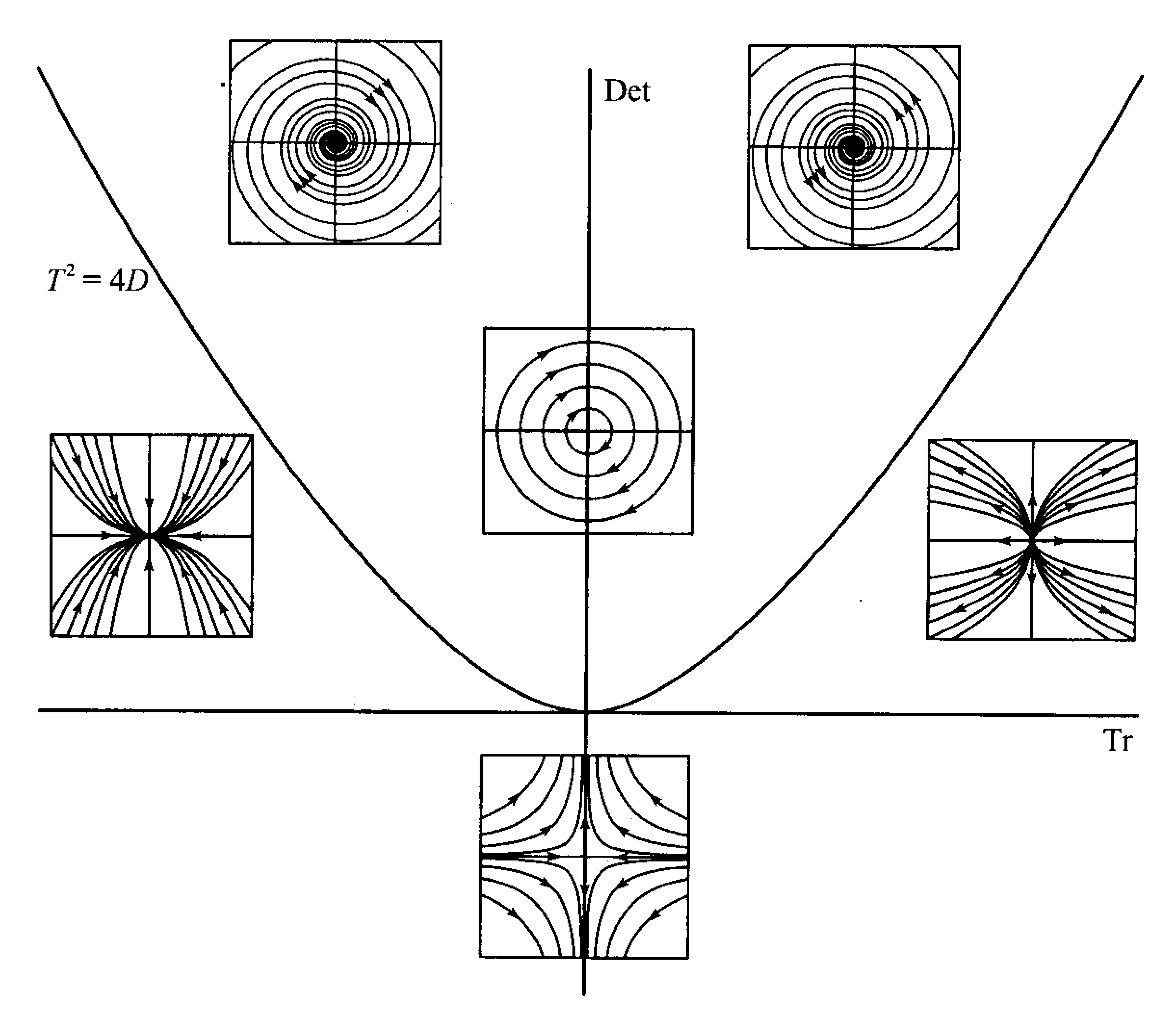


图 4.1 迹–行列式平面. 与作者面部的任何相似都纯属意外

最后需要说明的是, 即使不算出系统的特征值, 我们也可以从 D,T 的取值上得到系统相当多的信息, 例如, 只要 $D<0$, 我们就知道原点为一鞍点, 类似地, 如果 D 和 T 都是正的, 则原点是一源点.

4.2 动力学分类

在本节, 我们给出平面线性系统的另外一种更动力学的分类. 从动力系统的观点看, 我们通常关心微分方程解的长期行为. 如果两个系统的解在将来是一样的, 则它们就是等价的. 为了说得准确, 我们先回忆 1.5 节引入的一些概念.

为了强调解对时间和初值条件 $\boldsymbol{X}_0$ 的同时依赖性, 我们用 $\phi_t(X_0)$ 来表示满足初值条件 X_0 的解, 即 $\phi_0(X_0)=X_0$. 函数 $\phi(t,X_0)=\phi_t(X_0)$ 称为微分方程的**流**, 而 ϕ_t 则称为流的**时间** t **映射**.

例如, 对系统

$$\boldsymbol{X}'=\begin{pmatrix}2&0\\0&3\end{pmatrix}\boldsymbol{X},$$

其时间 t 映射就是

$$\phi_t(x_0, y_0) = \left(x_0 e^{2t}, y_0 e^{3t}\right).$$

由此可见, 流是同时依赖于时间和初值的函数.

我们认为两个系统是动力等价的, 如果存在函数 h, 它将一个流变为另一个流, 这里的函数要求为同胚, 即, h 是一对一的、满的连续函数, 并且它的逆也是连续的.

定义 假设 $\boldsymbol{X}' = \boldsymbol{AX}$ 和 $\boldsymbol{X}' = \boldsymbol{BX}$ 的流分别为 $\phi^{\boldsymbol{A}}$ 和 $\phi^{\boldsymbol{B}}$. 这两个系统称为 (拓扑)**共轭的**, 如果存在同胚 $h: \mathbb{R}^2 \to \mathbb{R}^2$ 满足

$$\phi^{\boldsymbol{B}}(t, h(X_0)) = h(\phi^{\boldsymbol{A}}(t, X_0)).$$

同胚 h 称为一个**共轭**. 因而共轭将 $\boldsymbol{X}' = \boldsymbol{AX}$ 的解曲线带到 $\boldsymbol{X}' = \boldsymbol{BX}$ 的解曲线. ■

例 对于两个一维系统

$$x' = \lambda_1 x \quad 和 \quad x' = \lambda_2 x,$$

它们的流分别为

$$\phi^j(t, x_0) = x_0 e^{\lambda_j t},$$

$j = 1, 2$. 假设 λ_1 和 λ_2 都是非零的并且有相同的符号. 令

$$h(x) = \begin{cases} x^{\lambda_2/\lambda_1} & 如果\ x \geqslant 0 \\ -|x|^{\lambda_2/\lambda_1} & 如果\ x < 0 \end{cases}$$

其中

$$x^{\lambda_2/\lambda_1} = \exp\left(\frac{\lambda_2}{\lambda_1}\ln(x)\right).$$

可见 h 为实直线的一个同胚. 我们断言 h 是 $x' = \lambda_1 x$ 和 $x' = \lambda_2 x$ 之间的一个共轭. 为此, 我们验证如下：当 $x_0 > 0$ 时,

$$h(\phi^1(t, x_0)) = \left(x_0 e^{\lambda_1 t}\right)^{\lambda_2/\lambda_1} = x_0^{\lambda_2/\lambda_1} e^{\lambda_2 t} = \phi^2(t, h(x_0)),$$

如所要证. 当 $x_0 < 0$ 时, 可以用同样的计算验证. ■

这里需要注意几点. 首先, λ_1 和 λ_2 必须有相同的符号, 因为否则的话就会有 $|h(0)| = \infty$, 此时 h 不可能是同胚. 这正好与动力学等价的概念相吻合：如果 λ_1 和 λ_2 有相同的符号, 它们的解要么同时趋于原点, 要么同时远离原点, 因而它们的解的行为是相似的. 此外, 如果 $\lambda_2 < \lambda_1$, 则 h 在原点不可微, 而如果 $\lambda_2 > \lambda_1$, 则 $h^{-1}(x) = x^{\lambda_1/\lambda_2}$ 在原点不可微. 这也正是我们要求 h 只是同胚而不是微分同胚

(具有可微逆的可微同胚) 的原因：如果假设了可微性，则必有 $\lambda_1 = \lambda_2$，这样得到的等价性概念不是很令人感兴趣.

这个例子给出了 (自治) 线性一阶微分方程的一种分类，这种分类与我们在第 1 章中的定性观察吻合. 这里只有三个共轭类：汇点、源点和一个特殊的“中间情形”，$x' = 0$，此时所有的解都是常数.

现在转到平面情形. 首先我们看到，此时只须确定矩阵为标准形所对应系统之间的共轭，因为在第 3 章已经看到，存在线性映射 $\boldsymbol{T}: \mathbb{R}^2 \to \mathbb{R}^2$ 将 $\boldsymbol{A}$ 化为标准形，而此时 $\boldsymbol{T}$ 将 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的流的时间 t 映射带到 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的时间 t 映射.

我们将要进行的平面线性系统分类与一维情形是类似的. 我们将不讨论系统的特征值实部为 0 的情形，这种情形的讨论留作习题.

定义 一个矩阵 $\boldsymbol{A}$ 称为**双曲的**，如果它的每一个特征值都具有非零实部. 此时我们也称系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 为双曲的. ■

定理 假设 2×2 矩阵 $\boldsymbol{A}_1$ 和 $\boldsymbol{A}_2$ 都是双曲的，则线性系统 $\boldsymbol{X}' = \boldsymbol{A}_i\boldsymbol{X}$ 为共轭的当且仅当这两个矩阵具有负实部的特征值的数目相同. ■

因而两个双曲矩阵所对应的线性系统是共轭的当且仅当它们的特征值集合都在下面的同一类中：

(1) 一个特征值为正而另一个负;

(2) 两个特征值都有负实部;

(3) 两个特征值都有正实部.

在证明定理之前，我们要注意到这个定理蕴涵，具有螺线汇点的系统与具有 (实) 汇点的系统是实共轭的. 这是当然的! 因为它们的相图虽然看起来很不一样，但是要看到这两个系统的所有解在将来都是相同的：在 $t \to \infty$ 时，它们都趋于原点.

证明 定理的必要性是显然的. 下面我们来证明定理的充分性.

由刚才的讨论，我们不妨假设所有的系统都是标准形. 证明分为以下三种情形.

情形 1

假设有两个线性系统 $\boldsymbol{X}' = \boldsymbol{A}_i\boldsymbol{X}, i = 1, 2$，每个 $\boldsymbol{A}_i$ 都有特征值 $\lambda_i < 0 < \mu_i$. 于是对每个系统，原点都是鞍点. 这是容易的情形. 前面已经看到，两个实微分方程 $x' = \lambda_i x$ 的流之间可以用同胚

$$h_1(x) = \begin{cases} x^{\lambda_2/\lambda_1} & \text{如果 } x \geqslant 0 \\ -|x|^{\lambda_2/\lambda_1} & \text{如果 } x < 0 \end{cases}$$

建立共轭. 类似地，方程 $y' = \mu_i y$ 的流之间可以用一个类似函数 h_2 建立共轭. 现

在, 定义

$$H(x,y) = (h_1(x), h_2(y)).$$

易见, H 就是两个系统之间的一个共轭.

情形 2

考虑系统 $\boldsymbol{X}' = \boldsymbol{AX}$, 其中 $\boldsymbol{A}$ 为标准形, 并且它的特征值的实部都是负的. 进一步, 我们假设矩阵 $\boldsymbol{A}$ 不是下面的形式

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

其中 $\lambda < 0$. 从而 $\boldsymbol{A}$ 为下面两个标准形之一:

$$\text{(a)} \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \qquad \text{(b)} \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix},$$

其中 $\alpha, \lambda, \mu < 0$. 我们将证明, 在每一种情形, 系统都与 $\boldsymbol{X}' = \boldsymbol{BX}$ 共轭, 其中

$$\boldsymbol{B} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

由此可得, 任何这种形式的两个系统都是共轭的.

考虑平面上的单位圆周, 将它看成一条参数曲线 $X(\theta) = (\cos\theta, \sin\theta)$, $0 \leqslant \theta \leqslant 2\pi$, 记此圆周为 S^1. 我们首先断言: 上面形式的矩阵所确定的向量场一定是指向 S^1 的内部. 在情形 (a), 向量场在 S^1 上的表达式为

$$\boldsymbol{AX}(\theta) = \begin{pmatrix} \alpha\cos\theta + \beta\sin\theta \\ -\beta\cos\theta + \alpha\sin\theta \end{pmatrix}.$$

而 S^1 在 $X(\theta)$ 处指向外面的法向量为

$$\boldsymbol{N}(\theta) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}.$$

由于 $\alpha < 0$, 这两个向量的点乘满足

$$\boldsymbol{AX}(\theta) \cdot \boldsymbol{N}(\theta) = \alpha(\cos^2\theta + \sin^2\theta) < 0.$$

这就说明了 $\boldsymbol{AX}(\theta)$ 的确指向 S^1 的内部. 情形 (b) 的验证更容易.

作为推论, 我们有 $\boldsymbol{X}' = \boldsymbol{AX}$ 的每个非零解正好穿过 S^1 一次. 用 $\phi_t^{\boldsymbol{A}}$ 记这个系统的时间 t 映射, 而用 $\tau = \tau(x,y)$ 记 $\phi_t^{\boldsymbol{A}}(x,y)$ 到达 S^1 的时刻. 于是有

$$\left|\phi_{\tau(x,y)}^{\boldsymbol{A}}(x,y)\right| = 1.$$

记 $\phi_t^{\boldsymbol{B}}$ 为系统 $\boldsymbol{X}'=\boldsymbol{BX}$ 的时间 t 映射, 显然,

$$\phi_t^{\boldsymbol{B}}(x,y)=(\mathrm{e}^{-t}x,\mathrm{e}^{-t}y).$$

现在来定义这两个系统之间的共轭 H. 如果 $(x,y)\neq(0,0)$, 令

$$H(x,y)=\phi_{-\tau(x,y)}^{\boldsymbol{B}}\phi_{\tau(x,y)}^{\boldsymbol{A}}(x,y),$$

再令 $H(0,0)=(0,0)$. 几何上, $H(x,y)$ 的取值按如下方式确定：先沿 $\boldsymbol{X}'=\boldsymbol{AX}$ 的解曲线 (向前或向后) 走 $\tau(x,y)$ 时间单位, 直到解到达 S^1 上的某点, 然后从该点出发, 沿着 $\boldsymbol{X}'=\boldsymbol{BX}$ 的解曲线反向走 τ 时间单位到达的点就是 $H(x,y)$, 见图 4.2.

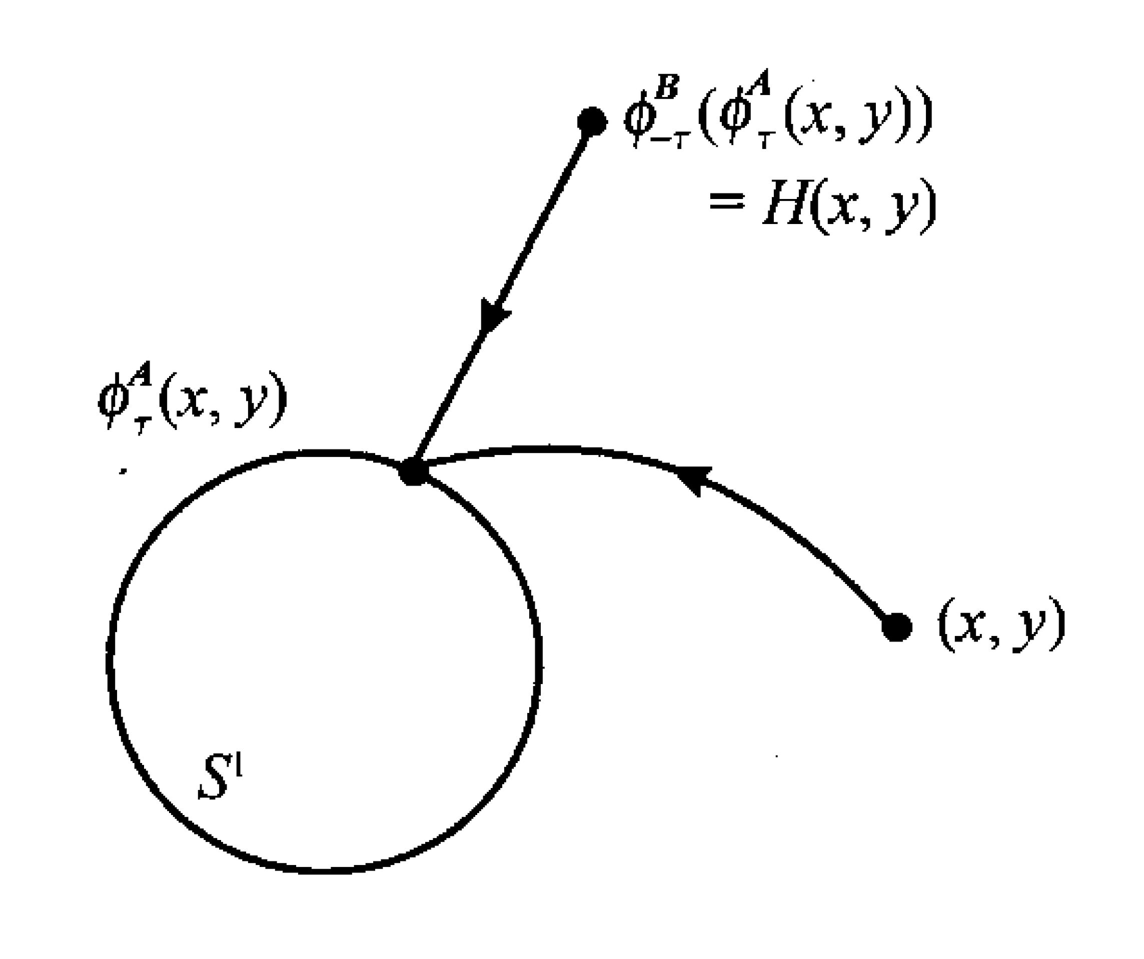


图4.2 $\tau(x,y)$ 的定义

为了看出 H 给出了一个共轭, 首先注意到

$$\tau(\phi_s^{\boldsymbol{A}}(x,y))=\tau(x,y)-s,$$

这是因为

$$\phi_{\tau-s}^{\boldsymbol{A}}\phi_s^{\boldsymbol{A}}(x,y)=\phi_\tau^{\boldsymbol{A}}(x,y)\in S^1.$$

于是我们就有

$$H\left(\phi_s^{\boldsymbol{A}}(x,y)\right)=\phi_{-\tau+s}^{\boldsymbol{B}}\phi_{\tau-s}^{\boldsymbol{A}}\left(\phi_s^{\boldsymbol{A}}(x,y)\right)=\phi_s^{\boldsymbol{B}}\phi_{-\tau}^{\boldsymbol{B}}\phi_\tau^{\boldsymbol{A}}(x,y)=\phi_s^{\boldsymbol{B}}\left(H(x,y)\right),$$

从而 H 为一共轭.

然后, 我们来说明 H 是一同胚. 简单地将定义 H 的过程反过来就可以定义出它的逆. 即, 令

$$G(x,y)=\phi_{-\tau_1(x,y)}^{\boldsymbol{A}}\phi_{\tau_1(x,y)}^{\boldsymbol{B}}(x,y),$$

及 $G(0,0)=(0,0)$, 这里 $\tau_1(x,y)$ 是 $\boldsymbol{X}'=\boldsymbol{BX}$ 过 (x,y) 的解到达 S^1 的时刻. 简单的计算表明 $\tau_1(x,y)=\ln r$, 其中 $r^2=x^2+y^2$. 显然有 $G=H^{-1}$, 从而 H 是一对一且满的. 又由于当 $(x,y)\neq(0,0)$ 时, G 可以改写成

$$G(x,y)=\phi_{-\ln r}^{\boldsymbol{A}}\left(\frac{x}{r},\frac{y}{r}\right),$$

这是一些连续函数的复合, 从而 G 在 (x,y) 点处连续. 再看 G 在原点的连续性. 假设 (x,y) 靠近原点, 于是 r 很小, 而且当 $r\to0$ 时, $-\ln r\to\infty$. 现在, $(x/r,y/r)$ 为 S^1 上一点, 而当 r 充分小时, $\phi_{-\ln r}^{\boldsymbol{A}}$ 将单位圆周映射到 $(0,0)$ 附近, 这就说明了 G 在 $(0,0)$ 处是连续的.

于是我们只须证明 H 的连续性. 为此, 需要说明 $\tau(x,y)$ 是连续的. τ 由方程

$$\left|\phi_t^{\boldsymbol{A}}(x,y)\right| = 1$$

所确定. 记 $\phi_t^{\boldsymbol{A}}(x,y) = (x(t), y(t))$. 通过求 $|\phi_t^{\boldsymbol{A}}(x,y)|$ 对 t 的偏导数可得

$$\begin{aligned}
\frac{\partial}{\partial t}\left|\phi_t^{\boldsymbol{A}}(x,y)\right| &= \frac{\partial}{\partial t}\sqrt{(x(t))^2+(y(t))^2} \\
&= \frac{1}{\sqrt{(x(t))^2+(y(t))^2}}\left(x(t)x'(t)+y(t)y'(t)\right) \\
&= \frac{1}{|\phi_t^{\boldsymbol{A}}(x,y)|}\left(\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}\cdot\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}\right).
\end{aligned}$$

在 $t=\tau(x,y)$ 时, 由于 $(x'(t), y'(t))$ 给出的向量场指向 S^1 的内部, 因而上面最后一式中的点乘是非零的, 于是在 $(\tau(x,y), x, y)$ 处有

$$\frac{\partial}{\partial t}\left|\phi_t^{\boldsymbol{A}}(x,y)\right| \neq 0.$$

由隐函数定理可知 τ 在 (x,y) 处是可微的, 从而也就连续. H 在原点处连续性的证明和 $G=H^{-1}$ 时类似. 于是 H 为同胚, 这样就得到了 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 和 $\boldsymbol{X}'=\boldsymbol{B}\boldsymbol{X}$ 之间的一个共轭.

可以看出, 这个证明在特征值都有正实部时一样可行.

情形 3

最后, 假设

$$\boldsymbol{A} = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},$$

$\lambda < 0$. 此时, 对应的向量场不一定指向单位圆的内部. 然而, 如果令

$$\boldsymbol{T} = \begin{pmatrix} 1 & 0 \\ 0 & \epsilon \end{pmatrix},$$

则只要 ϵ 足够小, 变换后的向量场

$$\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$$

的确是指向单位圆内部的. 事实上,

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda & \epsilon \\ 0 & \lambda \end{pmatrix},$$

从而

$$\left(T^{-1}AT\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\right)\cdot\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}=\lambda+\epsilon\sin\theta\cos\theta,$$

因而如果选取 $\epsilon<-\lambda$, 则上面的点乘为负. 这样, 经过变量替换 T 之后, 情形 2 的证明就可以适用了. 定理由此得证. ■

4.3 探索：一个 3D 参数空间

考虑线性系统的三参数族

$$X'=\begin{pmatrix}a & b\\ c & 0\end{pmatrix}X,$$

其中 a,b,c 为参数.

(1) 首先, 固定 $a>0$. 在 bc 平面上描述类似的迹-行列式平面, 即, 在 bc 平面上确定 b,c 的取值, 使得相应的系统具有鞍点、中心和螺线汇点等. 在 bc 平面上简略地描出这些区域.

(2) 当 $a<0$ 和 $a=0$ 时, 重复上面的问题.

(3) 描述当 a 从正变到负时出现的分岔.

(4) 然后将所有的这些信息综合起来, 给出系统的一个完整的三维参数空间的描述. 你可以建立这个空间的一个 3D 模型, 或者像翻书一样动画演示这些变化 (比如说, 当 a 变化时), 或者用计算机直观地画出图像. 不管用什么方式, 你的模型需要能准确地捕捉到所有不同区域的动力学行为.

习　题

1. 考虑线性系统的单参数族

$$X'=\begin{pmatrix}a & \sqrt{2}+(a/2)\\ \sqrt{2}-(a/2) & 0\end{pmatrix}X.$$

 (a) 简略地描出当 a 变化时, 这个线性系统族在迹-行列式平面上所走过的路线.

 (b) 讨论沿着这条路线所出现的分岔并算出对应的 a 值.

2. 对两参数族的系统

$$X'=\begin{pmatrix}a & b\\ b & a\end{pmatrix}X,$$

在 ab 平面上，简略地作出类似的迹-行列式平面，即，在 ab 平面上确定出一些区域，使得在每个区域上系统有相似的相图.

3. 考虑调和振子方程 (其中 $m=1$)

$$x''+bx'+kx=0,$$

其中 $b\geqslant 0,\ k>0$. 在 bk 平面的相应部分确定出一些区域，使得在每个区域上系统有相似的相图.

4. 证明 $H(x,y)=(x,-y)$ 是系统

$$\boldsymbol{X}'=\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\boldsymbol{X} \quad 和 \quad \boldsymbol{Y}'=\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\boldsymbol{Y}$$

之间的一个共轭.

5. 对下面的每一对系统，找出它们流之间一个共轭的显式表示.

(a) $\boldsymbol{X}'=\begin{pmatrix}-1 & 1\\ 0 & 2\end{pmatrix}\boldsymbol{X} \quad 和 \quad \boldsymbol{Y}'=\begin{pmatrix}1 & 0\\ 1 & -2\end{pmatrix}\boldsymbol{Y}$

(b) $\boldsymbol{X}'=\begin{pmatrix}0 & 1\\ -4 & 0\end{pmatrix}\boldsymbol{X} \quad 和 \quad \boldsymbol{Y}'=\begin{pmatrix}0 & 2\\ -2 & 0\end{pmatrix}\boldsymbol{Y}$.

6. 证明：如果两个线性系统具有相同的特征值 $\pm i\beta,\beta\neq 0$，则它们一定共轭. 如果两个线性系统的特征值分别为 $\pm i\beta$ 和 $\pm i\gamma$，且 $\beta\neq\gamma$，情况会怎么样？如果 $\gamma=-\beta$ 呢？

7. 考虑所有恰好有一个特征值为 0 的线性系统. 这些系统中哪些是共轭的？请给出证明.

8. 考虑所有具有两个 0 特征值的线性系统. 这些系统中哪些是共轭的？请给出证明.

9. 给出非双曲 2×2 系统共轭类的一个完整描述.

第 5 章　高维线性代数

与第 2 章一样, 在试图求解高维线性微分方程系统之前, 我们必须熟悉一下高维的线性代数. 虽然在高维, 矩阵的不同标准形个数变得多了, 但在作坐标变换将矩阵化为标准形的过程中所用到的代数思想大多都已在 2×2 情形时出现过了. 特别地, 当矩阵具有不同 (实的或复的) 特征值时, 除了增加很少的代数复杂性外, 可以作类似处理, 因而我们首先解决这种情形. 在 5.6 节, 我们会看到, 这是一种 "通有" 情形. 处理重特征值时需要用到更复杂的代数概念, 相应的背景知识将在 5.4 节给出. 在本章, 我们总是假定读者熟知如何将相应的矩阵化为 (简化) 行阶梯形, 进而求解线性代数方程系统.

5.1　线性代数预备知识

在本节, 我们要将 2.3 节中的许多代数概念推广到高维. 记一个向量 $X\in\mathbb{R}^n$ 的坐标形式为

$$\boldsymbol{X}=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}.$$

在平面上, 一对向量 $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 如果不共线就称为线性无关的, 等价地说, $\boldsymbol{V}$ 和 $\boldsymbol{W}$ 在称为线性无关的, 如果不存在 (非零的) 实数 α 和 β 使得 $\alpha\boldsymbol{V}+\beta\boldsymbol{W}$ 为零向量.

一般地, 在 $\mathbb{R}^n$ 中, 一组向量 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_k\in\mathbb{R}^n$ 称为**线性无关的**, 如果只要

$$\alpha_1\boldsymbol{V}_1+\cdots+\alpha_k\boldsymbol{V}_k=0$$

(其中 $\alpha_j\in\mathbb{R}$), 就有每一个 $\alpha_j=0$. 如果我们能找到不全为零的这样的 $\alpha_1,\cdots,\alpha_k$, 则称这组向量是**线性相关的**. 注意, 如果 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_k\in\mathbb{R}^n$ 是线性无关的, 而 $\boldsymbol{W}$ 是它们的一个线性组合

$$\boldsymbol{W}=\beta_1\boldsymbol{V}_1+\cdots+\beta_k\boldsymbol{V}_k,$$

则这样的 β_j 是唯一的, 这是因为, 如果

$$\boldsymbol{W}=\gamma_1\boldsymbol{V}_1+\cdots+\gamma_k\boldsymbol{V}_k,$$

则有

$$0 = \boldsymbol{W} - \boldsymbol{W} = (\beta_1 - \gamma_1)\boldsymbol{V}_1 + \cdots + (\beta_k - \gamma_k)\boldsymbol{V}_k,$$

由 $\boldsymbol{V}_j$ 的线性无关性, 则对所有的 j 必有 $\beta_j = \gamma_j$.

例　在 $\mathbb{R}^3$ 中, 向量组 (1,0,0),(0,1,0), (0,0,1) 显然是线性无关的. 更一般地, 令 E_j 是 $\mathbb{R}^n$ 中满足第 j 个分量是 1、其余分量都是 0 的向量. 则在 $\mathbb{R}^n$ 中, 向量组 $\boldsymbol{E}_1, \cdots, \boldsymbol{E}_n$ 是线性无关的. 向量组 $\boldsymbol{E}_1, \cdots, \boldsymbol{E}_n$ 的集合称为 $\mathbb{R}^n$ 的**标准基**. 我们将在 5.4 节讨论基的概念. ■

例　$\mathbb{R}^3$ 中的向量组 $(1,0,0),(1,1,0),(1,1,1)$ 是线性无关的, 这是因为, 如果有

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 + \alpha_3 \\ \alpha_2 + \alpha_3 \\ \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

则从第三个分量可得 $\alpha_3 = 0$, 将 $\alpha_3 = 0$ 代入第二个分量又可得 $\alpha_2 = 0$, 类似地再从第一个分量可得 $\alpha_1 = 0$. 另一方面, 向量组 $(1,1,1),(1,2,3),(2,3,4)$ 则是线性相关的, 因为我们有

$$1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - 1 \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$ ■

在求解线性微分方程系统时, 经常会遇到 $\mathbb{R}^n$ 的一类特殊子集合：子空间. $\mathbb{R}^n$ 的一个**子空间**就是一个给定的向量集合所有可能的线性组合的全体, 准确地说, 给定向量 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k \in \mathbb{R}^n$, 集合

$$\mathcal{S} = \{\alpha_1 \boldsymbol{V}_1 + \cdots + \alpha_k \boldsymbol{V}_k | \alpha_j \in \mathbb{R}\}$$

就是 $\mathbb{R}^n$ 的一个子空间. 此时, 我们称 $\mathcal{S}$ 是由 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k$**张成**的. 等价地, 可以证明 (见本章的习题 12), 如果 $\mathcal{S}$ 为 $\mathbb{R}^n$ 的一个非空子集, 则它为子空间当且仅当下面两条满足：

(1) 如果 $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{S}$, 则 $\boldsymbol{X} + \boldsymbol{Y} \in \mathcal{S}$;

(2) 如果 $\boldsymbol{X} \in \mathcal{S}, \alpha \in \mathbb{R}$, 则 $\alpha\boldsymbol{X} \in \mathcal{S}$.

注意, 零向量属于 $\mathbb{R}^n$ 的任何一个子空间, 而子空间 $\mathcal{S}$ 中向量的任一线性组合也属于 $\mathcal{S}$.

例　$\mathbb{R}^n$ 中过原点的任何直线都是 $\mathbb{R}^n$ 的一个子空间, 这是因为这样的直线可以写成 $\{t\boldsymbol{V} | t \in \mathbb{R}\}$, 其中 $\boldsymbol{V} \in \mathbb{R}^n$ 为一非零向量. 此时单个向量 $\boldsymbol{V}$ 就张成了这个子空间. 在 $\mathbb{R}^3$ 中由 $x + y + z = 0$ 确定的平面 $\mathcal{P}$ 是 $\mathbb{R}^3$ 的一个子空间, 事实上, $\mathcal{P}$ 中的

任何向量 $\boldsymbol{V}$ 都可以写成 $(x, y, -x-y)$ 的形式, 也即

$$\boldsymbol{V} = x\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

这也表明向量 $(1, 0, -1)$ 和 $(0, 1, -1)$ 张成 $\mathcal{P}$. ■

在线性代数中, 经常会遇到长方形的 $n \times m$ 矩阵, 但在微分方程中, 最常遇到的是 $(n \times n)$ 的方阵, 因而在本章我们总假设矩阵都是 $n \times n$ 的. 我们将这样的一个矩阵写成

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

或写得紧凑些, $\boldsymbol{A} = [a_{ij}]$.

设 $\boldsymbol{X} = (x_1, \cdots, x_n) \in \mathbb{R}^n$, 定义乘积 $\boldsymbol{AX}$ 为向量

$$\boldsymbol{AX} = \begin{pmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{pmatrix},$$

其中这个向量的第 i 个元素为 $\boldsymbol{A}$ 的第 i 行与向量 $\boldsymbol{X}$ 的点乘.

矩阵的和可以按自然的方式定义: 设 $\boldsymbol{A} = [a_{ij}], \boldsymbol{B} = [b_{ij}]$ 为两个 $n \times n$ 矩阵, 则定义 $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C}$, 其中 $\boldsymbol{C} = [a_{ij} + b_{ij}]$. 矩阵运算具有如下明显的线性性质:

(1) $\boldsymbol{A}(k_1\boldsymbol{X}_1 + k_2\boldsymbol{X}_2) = k_1\boldsymbol{AX}_1 + k_2\boldsymbol{AX}_2$, 其中 $k_j \in \mathbb{R}, \boldsymbol{X}_j \in \mathbb{R}^n$;

(2) $\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$;

(3) $(\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C} = \boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C})$.

设 $\boldsymbol{A}, \boldsymbol{B}$ 为两个 $n \times n$ 矩阵, 则它们的乘积定义为 $n \times n$ 矩阵 $\boldsymbol{AB} = [c_{ij}]$, 其中

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj},$$

即 c_{ij} 是 $\boldsymbol{A}$ 的第 i 行与 $\boldsymbol{B}$ 的第 j 列的点乘. 容易验证, 如果 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ 都是 $n \times n$ 矩阵, 则有

(1) $(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC})$;

(2) $\boldsymbol{A}(\boldsymbol{B}+\boldsymbol{C})=\boldsymbol{AB}+\boldsymbol{AC}$;

(3) $(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{C}=\boldsymbol{AC}+\boldsymbol{BC}$;

(4) $k(\boldsymbol{AB})=(k\boldsymbol{A})\boldsymbol{B}=\boldsymbol{A}(k\boldsymbol{B})$, 其中 $k\in\mathbb{R}$.

所有上面有关矩阵运算的性质都可以很容易验证, 只须写出相应矩阵的 ij 元素就可以了. 需要牢记的是矩阵乘法不可交换, 即, 通常 $\boldsymbol{AB}\neq\boldsymbol{BA}$. 例如,

$$\begin{pmatrix}1&0\\1&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}=\begin{pmatrix}1&1\\1&2\end{pmatrix},$$

然而

$$\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\1&1\end{pmatrix}=\begin{pmatrix}2&1\\1&1\end{pmatrix}.$$

此外, 矩阵消去律通常也是不成立的; 即, 我们不能从 $\boldsymbol{AB}=\boldsymbol{AC}$ 推出 $\boldsymbol{B}=\boldsymbol{C}$, 例如:

$$\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}1&0\\0&0\end{pmatrix}=\begin{pmatrix}1&0\\1&0\end{pmatrix}=\begin{pmatrix}1&1\\1&1\end{pmatrix}\begin{pmatrix}1/2&1/2\\1/2&-1/2\end{pmatrix}.$$

特别地, 如果 $\boldsymbol{AB}$ 为零矩阵, 并不能得出 $\boldsymbol{A},\boldsymbol{B}$ 中有一个为零矩阵.

$n\times n$ 矩阵 $\boldsymbol{A}$ 称为**可逆的**, 如果存在 $n\times n$ 矩阵 $\boldsymbol{C}$ 使得 $\boldsymbol{AC}=\boldsymbol{CA}=\boldsymbol{I}$, 其中 $\boldsymbol{I}$ 为 $n\times n$ 单位矩阵, 它的元素沿对角线取 1, 而其他地方为 0. $\boldsymbol{C}$ 称为 $\boldsymbol{A}$ 的**逆**. 注意, 如果 $\boldsymbol{A}$ 有逆, 则逆是唯一的, 这是因为如果还有 $\boldsymbol{AB}=\boldsymbol{BA}=\boldsymbol{I}$, 则

$$\boldsymbol{C}=\boldsymbol{CI}=\boldsymbol{C}(\boldsymbol{AB})=(\boldsymbol{CA})\boldsymbol{B}=\boldsymbol{IB}=\boldsymbol{B}.$$

$\boldsymbol{A}$ 的逆记为 $\boldsymbol{A}^{-1}$.

如果 $\boldsymbol{A}$ 可逆, 则对任意的 $\boldsymbol{V}\in\mathbb{R}^n$, 向量方程 $\boldsymbol{AX}=\boldsymbol{V}$ 都有唯一解, 事实上 $\boldsymbol{A}^{-1}\boldsymbol{V}$ 就是一个解, 而且只有这个解, 这是因为如果 $\boldsymbol{Y}$ 是另一个解, 则有

$$\boldsymbol{Y}=(\boldsymbol{A}^{-1}\boldsymbol{A})\boldsymbol{Y}=\boldsymbol{A}^{-1}(\boldsymbol{AY})=\boldsymbol{A}^{-1}\boldsymbol{V}.$$

至于这个断言的逆命题, 回忆一下, 方程 $\boldsymbol{AX}=\boldsymbol{V}$ 有唯一解当且仅当矩阵 $\boldsymbol{A}$ 的行简化阶梯形为单位矩阵. $\boldsymbol{A}$ 的行**简化阶梯形**是通过一系列的**初等行变换**得来的, 而初等行变换是指下面的三种变换:

(1) 将 $\boldsymbol{A}$ 的第 i 行的 k 倍加到第 j 行; (2) 交换第 i 行和第 j 行; (3) 将第 i 行乘上 $k\neq0$.

注意这些初等行变换正好对应于通常求解线性代数方程系统的以下变换:

(1) 将第 i 个方程的 k 倍加到第 j 个方程;

(2) 交换第 i 个方程和第 j 个方程;

(3) 将第 i 个方程乘上 $k\neq0$.

每一个初等行变换都可以用 $\boldsymbol{A}$ 乘上一个**初等矩阵**来表示. 例如, 如果记矩阵 $\boldsymbol{L} = [l_{ij}]$ 是这样的矩阵, 它沿对角线取值 1, 而对某给定的 $i, j, i \neq j$, $l_{ji} = k$, 其他元素都为 0, 则 $\boldsymbol{LA}$ 就是 $\boldsymbol{A}$ 通过第 1 类行变换得到的矩阵. 同样地, 如果 $l_{ii} = l_{jj} = 0, l_{ij} = l_{ji} = 1$, 且 $\boldsymbol{L}$ 在对角线上的其他位置取值为 1, 而且所有其他位置取 0, 则 $\boldsymbol{LA}$ 就是 $\boldsymbol{A}$ 通过第 2 类行变换得到的矩阵. 最后, 如果 $\boldsymbol{L}$ 是将单位矩阵中的 ii 位置的 1 换成 k 得到的, 则 $\boldsymbol{LA}$ 就是 $\boldsymbol{A}$ 通过第 3 类行变换得到的矩阵. 上述三种形式的每一个矩阵 $\boldsymbol{L}$ 都称为一个初等矩阵.

每一个初等矩阵都是可逆的, 这是因为将对应的初等行变换逆向操作就可以得到它的逆. 作为推论, 初等矩阵的乘积都是可逆的. 于是, 如果将对 $\boldsymbol{A}$ 作初等行变换化为行简化阶梯形 (为单位矩阵) 的对应的初等矩阵记为 $\boldsymbol{L}_1, \cdots, \boldsymbol{L}_n$, 则有 $(\boldsymbol{L}_n \cdot \cdots \cdot \boldsymbol{L}_1) = \boldsymbol{A}^{-1}$. 即, 如果对任意的 $\boldsymbol{V} \in \mathbb{R}^n$, 向量方程 $\boldsymbol{AX} = \boldsymbol{V}$ 有唯一解, 则 $\boldsymbol{A}$ 可逆. 这样就得到了我们的第一个重要结果.

命题 设 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵, 则代数方程系统 $\boldsymbol{AX} = \boldsymbol{V}$ 对任意 $\boldsymbol{V} \in \mathbb{R}^n$ 都有唯一解当且仅当 $\boldsymbol{A}$ 可逆. ■

现在自然就要问: 如何判定 $\boldsymbol{A}$ 是否可逆? 下面的命题给出了此问题的一种回答.

命题 矩阵 $\boldsymbol{A}$ 可逆当且仅当 $\boldsymbol{A}$ 的列向量组构成一个线性无关的向量组.

证明 先假设 $\boldsymbol{A}$ 可逆, 其列向量为 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_n$. 于是我们有 $\boldsymbol{AE}_j = \boldsymbol{V}_j$, 其中 $\boldsymbol{E}_j$ 为 $\mathbb{R}^n$ 的标准基. 如果 $\boldsymbol{V}_j$ 不是线性无关的, 则可以找到不全为零的实数 $\alpha_1, \cdots, \alpha_n$ 使得 $\sum_j \alpha_j \boldsymbol{V}_j = 0$, 但是

$$0 = \sum_{j=1}^{n} \alpha_j \boldsymbol{AE}_j = \boldsymbol{A} \left(\sum_{j=1}^{n} \alpha_j \boldsymbol{E}_j \right).$$

于是方程 $\boldsymbol{AX} = \boldsymbol{0}$ 就有了两个解, 一个是非零向量 $(\alpha_1, \cdots, \alpha_n)$, 另一个是 $\boldsymbol{0}$ 向量, 这与上一个命题矛盾.

反过来, 假设 $\boldsymbol{V}_j$ 是线性无关的. 如果 $\boldsymbol{A}$ 不可逆, 则可以找到一对向量 $\boldsymbol{X}_1, \boldsymbol{X}_2$, $\boldsymbol{X}_1 \neq \boldsymbol{X}_2$, 使得 $\boldsymbol{AX}_1 = \boldsymbol{AX}_2$, 从而非零向量 $\boldsymbol{Z} = \boldsymbol{X}_1 - \boldsymbol{X}_2$ 就满足 $\boldsymbol{AZ} = \boldsymbol{0}$. 记 $\boldsymbol{Z} = (\alpha_1, \cdots, \alpha_n)$, 于是有

$$\boldsymbol{0} = \boldsymbol{AZ} = \sum_{j=1}^{n} \alpha_j \boldsymbol{V}_j,$$

这说明 $\boldsymbol{V}_j$ 不是线性无关的, 这一矛盾就证明了命题. ■

与 2×2 情形一样, 另一个判定矩阵是否可逆的判据是利用行列式, 它更加具有可计算性. 对于给定的矩阵 $\boldsymbol{A}$, 记 $\boldsymbol{A}_{ij}$ 为删掉 $\boldsymbol{A}$ 的第 i 行和第 j 列后得到的 $(n-1) \times (n-1)$ 矩阵.

定义　$\boldsymbol{A}=[a_{ij}]$ 的**行列式**可由下式归纳地定义：

$$\det \boldsymbol{A}=\sum_{k=1}^{n}(-1)^{1+k}a_{1k}\det \boldsymbol{A}_{1k}.$$

注意, 我们已经知道 2×2 矩阵的行列式, 因而上面的归纳法在 $k>2$ 时有意义. ■

例　根据定义, 我们来计算

$$\begin{aligned}\det\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} &= 1\det\begin{pmatrix}5&6\\8&9\end{pmatrix}-2\det\begin{pmatrix}4&6\\7&9\end{pmatrix}+3\det\begin{pmatrix}4&5\\7&8\end{pmatrix}\\ &= -3+12-9=0.\end{aligned}$$

需要说明的是, 上面 $\det \boldsymbol{A}$ 的定义是按 $\boldsymbol{A}$ 的第一行展开的, 同样, 我们还可以按第 j 行展开：

$$\det \boldsymbol{A}=\sum_{k=1}^{n}(-1)^{j+k}a_{jk}\det \boldsymbol{A}_{jk}.$$

我们不打算证明这件事, 其证明无非就是一些直接的计算, 但很繁琐. 同样, $\det \boldsymbol{A}$ 还可以通过按列展开来计算 (见习题 1). ■

例　将上例中的矩阵按第二行和第三行展开可以得到同样的结果, 计算如下：

$$\begin{aligned}\det\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} &= -4\det\begin{pmatrix}2&3\\8&9\end{pmatrix}+5\det\begin{pmatrix}1&3\\7&9\end{pmatrix}-6\det\begin{pmatrix}1&2\\7&8\end{pmatrix}\\ &= 24-60+36=0\\ &= 7\det\begin{pmatrix}2&3\\5&6\end{pmatrix}-8\det\begin{pmatrix}1&3\\4&6\end{pmatrix}+9\det\begin{pmatrix}1&2\\4&5\end{pmatrix}\\ &= -21+48-27=0.\end{aligned}$$

同时, 请注意这个矩阵不可逆, 这是因为

$$\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix}\begin{pmatrix}1\\-2\\1\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}.$$ ■

某些特殊类型的矩阵的行列式是很容易计算的. 一个矩阵 $[a_{ij}]$ 称为上三角的, 如果主对角下面的元素都是 0, 即, 当 $i>j$ 时, $a_{ij}=0$. 可以类似地定义下三角矩阵. 我们有

命题 如果 $\boldsymbol{A}$ 是一个 $n\times n$ 上 (下) 三角矩阵, 则 $\det \boldsymbol{A}$ 就是对角线上所有元素的乘积, 即 $\det[a_{ij}] = a_{11}\cdots a_{nn}$. ■

证明可运用归纳法直接得到. 下面的命题描述了初等行变换对矩阵行列式的影响.

命题 设 $\boldsymbol{A},\boldsymbol{B}$ 为两个 $n\times n$ 矩阵.

(1) 如果 $\boldsymbol{B}$ 是由 $\boldsymbol{A}$ 的一行的某个倍数加到另一行得到的, 则 $\det \boldsymbol{B} = \det \boldsymbol{A}$.

(2) 如果 $\boldsymbol{B}$ 是由交换 $\boldsymbol{A}$ 的两行得到的, 则 $\det \boldsymbol{B} = -\det \boldsymbol{A}$.

(3) 如果 $\boldsymbol{B}$ 是由 $\boldsymbol{A}$ 的一行每个元素乘以 k 得到的, 则 $\det \boldsymbol{B} = k\det \boldsymbol{A}$.

证明 当 $\boldsymbol{A}$ 为一 2×2 矩阵时, 命题的证明是直接的, 下面我们用归纳法来证明. 假设 $\boldsymbol{A}$ 为 $k\times k$ 的, $k>2$. 为了计算 $\det \boldsymbol{B}$, 按一未被行变换影响的行来展开. 由行列式的归纳定义, 可见 $\det \boldsymbol{B}$ 是一些 $(k-1)\times(k-1)$ 大小的行列式的和. 而且这些子行列式都是 $\boldsymbol{A}$ 的相应的子行列式经过与 $\boldsymbol{A}$ 相同的行变换得到的. 由归纳法, 这些子行列式在情形 1, 2, 3, 将是 $\boldsymbol{A}$ 的相应的子行列式分别乘以 $1,-1,k$, 从而 $\det \boldsymbol{B}$ 也是 $\det \boldsymbol{A}$ 分别乘以 $1,-1,k$. ■

特别地, 我们还应注意到, 当 L 是一个初等矩阵时, 则有

$$\det(\boldsymbol{LA}) = (\det \boldsymbol{L})(\det \boldsymbol{A}).$$

事实上, 对于上面的情形 1, 2, 3, $\det \boldsymbol{L} = 1,-1,k$(见习题 7). 由上面的命题, 我们就得到了 $\boldsymbol{A}$ 是否可逆的一个判据:

推论 (**可逆性判据**) 矩阵 $\boldsymbol{A}$ 可逆当且仅当 $\det \boldsymbol{A}\neq 0$.

证明 我们可将任何矩阵 $\boldsymbol{A}$ 通过初等行变换化为一个上三角矩阵 $\boldsymbol{B}$. 于是, $\boldsymbol{A}$ 可逆当且仅当这一行简化阶梯形矩阵 $\boldsymbol{B}$ 的对角线元素都非零. 特别地, 这一矩阵的行列式非零. 现在, 根据刚才的观察, $\boldsymbol{A}$ 进行行变化后的矩阵 $\boldsymbol{B}$ 的行列式就是 det $\boldsymbol{A}$ 乘上一个非零数, 由此可见, $\boldsymbol{B}$ 的对角线上的元素都非零当且仅当 $\det \boldsymbol{A}$ 非零. 证毕. ■

我们用行列式的另外一个重要性质来结束本节.

命题 $\det(\boldsymbol{AB}) = (\det \boldsymbol{A})(\det \boldsymbol{B})$.

证明 如果 $\boldsymbol{A},\boldsymbol{B}$ 中有一个是不可逆的, 则 $\boldsymbol{AB}$ 也是不可逆的 (见习题 11), 于是等式两端都是零, 从而命题成立. 下面假设 $\boldsymbol{A}$ 可逆, 则 $\boldsymbol{A}$ 可以写成初等矩阵的乘积:

$$\boldsymbol{A} = \boldsymbol{L}_1\cdots\boldsymbol{L}_n\cdot\boldsymbol{I}.$$

从而有

$$\begin{aligned}\det\,(\boldsymbol{AB}) &= \det\,(\boldsymbol{L}_1\cdots\boldsymbol{L}_n\boldsymbol{B}) = \det\,(\boldsymbol{L}_1)\det\,(\boldsymbol{L}_2\cdots\boldsymbol{L}_n\boldsymbol{B})\\ &= \det\,(\boldsymbol{L}_1)(\det \boldsymbol{L}_2)\cdots(\det \boldsymbol{L}_n)(\det \boldsymbol{B})\\ &= \det\,(\boldsymbol{L}_1\cdots\boldsymbol{L}_n)\det\,(\boldsymbol{B}) = \det\,(\boldsymbol{A})\det(\boldsymbol{B}).\end{aligned}$$ ■

5.2　特征值和特征向量

在第 3 章我们已经看到, 特征值和特征向量在求解线性微分方程系统的过程中起到了至关重要的作用.

定义　一个向量 $\boldsymbol{V}$ 称为 $n \times n$ 矩阵 $\boldsymbol{A}$ 的一个**特征向量**, 如果 $\boldsymbol{V}$ 是线性方程组系统 $(\boldsymbol{A}-\lambda \boldsymbol{I})\boldsymbol{V}=\boldsymbol{0}$ 的一个非零解. 其中的数量 λ 称为 $\boldsymbol{A}$ 的一个**特征值**, 而称 $\boldsymbol{V}$ 是属于 (相应于)λ 的一个特征向量. ■

在第 2 章中已经看到, 矩阵 $\boldsymbol{A}$ 的特征值可能是实的, 也可能是复的, 而属于它的特征向量也可能会有复分量.

由上一节的可逆性判据, λ 为 $\boldsymbol{A}$ 的特征值当且仅当 λ 是**特征方程**

$$\det(\boldsymbol{A}-\lambda \boldsymbol{I})=0$$

的一个根. 由于 $\boldsymbol{A}$ 是 $n \times n$ 的, 特征方程是一个 n 次多项式方程, 因而正好有 n 个根 (计重数).

在 $\mathbb{R}^2$ 时, 我们已经看到, 微分方程系统的解有很多不同类型, 这些类型依赖于 $\boldsymbol{A}$ 的特征值的结构以及相应的标准形. 在高维, 将有更多类型的标准形, 我们将在本节及随后的几节中描述这些类型, 但我们将把一些事实的证明留作习题.

首先假设 $\lambda_1, \cdots, \lambda_l$ 是 $\boldsymbol{A}$ 的实不同特征值, $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_l$ 为分别属于它们的特征向量, 这里 “不同” 指的是没有两个特征值相同. 于是, 对每一个 k, $\boldsymbol{A}\boldsymbol{V}_k=\lambda_k \boldsymbol{V}_k$. 我们断言 $\boldsymbol{V}_k$ 为线性无关的. 否则, 取 $\boldsymbol{V}_k$ 的一个最大线性无关子集, 不妨设为 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_j$. 于是其他的特征向量以唯一的方式写成 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_j$ 的线性组合. 比如, $\boldsymbol{V}_{j+1}$ 就是的, 因而可以找到不全为 0 的 α_i 使得

$$\boldsymbol{V}_{j+1}=\alpha_1 \boldsymbol{V}_1+\cdots+\alpha_j \boldsymbol{V}_j.$$

在这个方程两端用 $\boldsymbol{A}$ 作用后可得

$$\lambda_{j+1} \boldsymbol{V}_{j+1}=\alpha_1 \boldsymbol{A}\boldsymbol{V}_1+\cdots+\alpha_j \boldsymbol{A}\boldsymbol{V}_j=\alpha_1 \lambda_1 \boldsymbol{V}_1+\cdots+\alpha_j \lambda_j \boldsymbol{V}_j.$$

此时有 $\lambda_{j+1} \neq 0$, 因为否则的话, 我们就有

$$\alpha_1 \lambda_1 \boldsymbol{V}_1+\cdots+\alpha_j \lambda_j \boldsymbol{V}_j=0,$$

且式子中的每一个 $\lambda_i \neq 0$, 这与 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_j$ 是线性无关的矛盾. 于是, 我们得到

$$\boldsymbol{V}_{j+1}=\alpha_1 \frac{\lambda_1}{\lambda_{j+1}} \boldsymbol{V}_1+\cdots+\alpha_j \frac{\lambda_j}{\lambda_{j+1}} \boldsymbol{V}_j.$$

由于 λ_i 互不相同, 这样 $\boldsymbol{V}_{j+1}$ 就可以有两种不同的方式写成 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_j$ 的线性组合, 这与这个向量组的线性无关性矛盾. 这样我们就证明了:

命题 假设 $\lambda_1,\cdots,\lambda_l$ 是 $\boldsymbol{A}$ 的实不同特征值, $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_l$ 为分别属于它们的特征向量, 则 $\boldsymbol{V}_j$ 是线性无关的. ■

当我们回头考虑微分方程时, 重要的则是:

推论 假设 $\boldsymbol{A}$ 为 $n\times n$ 矩阵, 具有实不同特征值, 则存在矩阵 $\boldsymbol{T}$ 使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},$$

其中对角线以外的元素都是 0.

证明 设 $\boldsymbol{V}_j$ 是属于 λ_j 的特征向量. 定义线性映射 $\boldsymbol{T}$ 使得 $\boldsymbol{T}\boldsymbol{E}_j=\boldsymbol{V}_j$, 其中 $\boldsymbol{E}_j$ 是 $\mathbb{R}^n$ 的标准基. 即, $\boldsymbol{T}$ 是以 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_n$ 为列向量的矩阵. 因为 $\boldsymbol{V}_j$ 是线性无关的, 故 $\boldsymbol{T}$ 可逆, 从而就有

$$(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{E}_j=\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{V}_j=\lambda_j\boldsymbol{T}^{-1}\boldsymbol{V}_j=\lambda_j\boldsymbol{E}_j.$$

即, $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 的第 j 列正好就是向量 $\lambda_j\boldsymbol{E}_j$, 这也就是我们要证的. ■

例 令

$$\boldsymbol{A}=\begin{pmatrix}1 & 2 & -1\\ 0 & 3 & -2\\ 0 & 2 & -2\end{pmatrix},$$

将 $\det(\boldsymbol{A}-\lambda\boldsymbol{I})$ 按第一列展开, 可得 $\boldsymbol{A}$ 的特征多项式为

$$\begin{aligned}\det(\boldsymbol{A}-\lambda\boldsymbol{I}) &= (1-\lambda)\det\begin{pmatrix}3-\lambda & -2\\ 2 & -2-\lambda\end{pmatrix}\\ &= (1-\lambda)((3-\lambda)(-2-\lambda)+4)\\ &= (1-\lambda)(\lambda-2)(\lambda+1),\end{aligned}$$

于是特征值为 $2,1,-1$. 属于 $\lambda=2$ 的特征向量可以通过求解方程组 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{X}=\boldsymbol{0}$ 得到, 于是就有

$$\begin{aligned}-x+2y-z &= 0\\ y-2z &= 0\\ 2y-4z &= 0,\end{aligned}$$

这个方程组可以化简为

$$\begin{aligned}x-3z &= 0\\ y-2z &= 0,\end{aligned}$$

从而 $\boldsymbol{V}_1=(3,2,1)$ 就是属于 $\lambda=2$ 的一个特征向量. 用同样的方式, 我们可以找到 $(1,0,0)$ 是属于 $\lambda=1$ 的一个特征向量, 而 $(0,1,2)$ 是属于 $\lambda=-1$ 的一个特征向量. 然后我们令

$$\boldsymbol{T}=\begin{pmatrix}3&1&0\\2&0&1\\1&0&2\end{pmatrix}.$$

简单的计算表明

$$\boldsymbol{A}\boldsymbol{T}=\boldsymbol{T}\begin{pmatrix}2&0&0\\0&1&0\\0&0&-1\end{pmatrix}.$$

由于 $\det\boldsymbol{T}=-3$, $\boldsymbol{T}$ 可逆, 从而我们就有

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}2&0&0\\0&1&0\\0&0&-1\end{pmatrix}.$$ ■

5.3　复特征值

现在我们来处理 $\boldsymbol{A}$ 具有非实 (复) 特征值的情形. 假设 $\alpha+\mathrm{i}\beta$ 是 $\boldsymbol{A}$ 的一个特征值, 其中 $\beta\neq0$. 由于 $\boldsymbol{A}$ 的特征方程具有实系数, 因而, 如果 $\alpha+\mathrm{i}\beta$ 是一个特征值, 则它的复共轭 $\overline{\alpha+\mathrm{i}\beta}=\alpha-\mathrm{i}\beta$ 也是一个特征值.

看出这点的另外一个方法如下：设 $\boldsymbol{V}$ 是属于 $\alpha+\mathrm{i}\beta$ 的一个特征向量, 于是由方程

$$\boldsymbol{A}\boldsymbol{V}=(\alpha+\mathrm{i}\beta)\boldsymbol{V}$$

可知 $\boldsymbol{V}$ 是一个具有复分量的向量. 我们记

$$\boldsymbol{V}=\begin{pmatrix}x_1+\mathrm{i}y_1\\\vdots\\x_n+\mathrm{i}y_n\end{pmatrix}.$$

用 $\overline{\boldsymbol{V}}$ 表示 $\boldsymbol{V}$ 的复共轭：

$$\overline{\boldsymbol{V}}=\begin{pmatrix}x_1-\mathrm{i}y_1\\\vdots\\x_n-\mathrm{i}y_n\end{pmatrix}.$$

于是我们就有

$$A\overline{V} = \overline{AV} = (\overline{\alpha + \mathrm{i}\beta})\overline{V} = (\alpha - \mathrm{i}\beta)\overline{V},$$

这表明 $\overline{V}$ 是属于特征值 $\alpha - \mathrm{i}\beta$ 的一个特征向量.

注意, 我们已经 (暂时地) 从 $\mathbb{R}^n$ 的 "实" 世界进入到了复向量的世界 $\mathbb{C}^n$. 这点不会真的产生什么困难, 理由是: 以前线性代数的结论在复向量时一样成立.

现在假设 A 是一个 $2n \times 2n$ 矩阵, 它具有不同的非实特征值 $\alpha_j \pm \mathrm{i}\beta_j$, $j = 1, \cdots, n$, 记 $V_j, \overline{V_j}$ 为属于它们的特征向量. 那么与上节的命题一样, 所有这些特征向量构成的向量组是线性无关的, 即, 如果有

$$\sum_{j=1}^{n}(c_j V_j + d_j \overline{V_j}) = 0,$$

其中 c_j, d_j 现在为复数, 则对所有的 j, 必有 $c_j = d_j = 0$.

现在我们来做坐标变换将 A 化为标准形. 令

$$W_{2j-1} = \frac{1}{2}(V_j + \overline{V_j})$$

$$W_{2j} = \frac{-\mathrm{i}}{2}(V_j - \overline{V_j}).$$

注意这里 W_{2j-1}, W_{2j} 都是实向量, 事实上, W_{2j-1} 是 V_j 的实部, 而 W_{2j} 是 V_j 的虚部. 利用 W_j 进行操作就把我们又带回 $\mathbb{R}^n$ 的世界.

命题 向量组 $W_1, \cdots, W_{2n}$ 是线性无关的.

证明 否则, 可以找到不全为零的实数 c_j, d_j, $j = 1, \cdots, n$ 使得

$$\sum_{j=1}^{n}\left(c_j W_{2j-1} + d_j W_{2j}\right) = 0.$$

于是就有

$$\frac{1}{2}\sum_{j=1}^{n}\left(c_j(V_j + \overline{V_j}) - \mathrm{i}d_j(V_j - \overline{V_j})\right) = 0,$$

由此可得

$$\sum_{j=1}^{n}\left((c_j - \mathrm{i}d_j)V_j + (c_j + \mathrm{i}d_j)\overline{V_j}\right) = 0.$$

由于 $V_j, \overline{V_j}$ 构成的向量组是线性无关的, 我们必有 $c_j \pm \mathrm{i}d_j = 0$, 由此可得, 对所有的 j, $c_j = d_j = 0$. 这一矛盾就证明了结论. ∎

注意, 此时我们有

$$\begin{aligned} \boldsymbol{A}\boldsymbol{W}_{2j-1} &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{V}_j + \boldsymbol{A}\overline{\boldsymbol{V}_j}) \\ &= \frac{1}{2}\big((\alpha + \mathrm{i}\beta)\boldsymbol{V}_j + (\alpha - \mathrm{i}\beta)\overline{\boldsymbol{V}_j}\big) \\ &= \frac{\alpha}{2}(\boldsymbol{V}_j + \overline{\boldsymbol{V}_j}) + \frac{\mathrm{i}\beta}{2}(\boldsymbol{V}_j - \overline{\boldsymbol{V}_j}) \\ &= \alpha\boldsymbol{W}_{2j-1} - \beta\boldsymbol{W}_{2j}. \end{aligned}$$

同理可以算得

$$\boldsymbol{A}\boldsymbol{W}_{2j} = \beta\boldsymbol{W}_{2j-1} + \alpha\boldsymbol{W}_{2j}.$$

现在定义线性映射 $\boldsymbol{T}$ 使得 $\boldsymbol{T}\boldsymbol{E}_j = \boldsymbol{W}_j, j = 1, \cdots, 2n$, 即, 与 $\boldsymbol{T}$ 相关的矩阵的列向量为 $\boldsymbol{W}_1, \cdots, \boldsymbol{W}_{2n}$. 可见, 这个矩阵具有实元素. 由于向量组 $\boldsymbol{W}_j$ 是线性无关的, 从 5.1 节可知, $\boldsymbol{T}$ 可逆. 现在来看矩阵 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$, 我们有

$$(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{E}_{2j-1} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{W}_{2j-1} = \boldsymbol{T}^{-1}(\alpha\boldsymbol{W}_{2j-1} - \beta\boldsymbol{W}_{2j}) = \alpha\boldsymbol{E}_{2j-1} - \beta\boldsymbol{E}_{2j},$$

类似地

$$(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{E}_{2j} = \beta\boldsymbol{E}_{2j-1} + \alpha\boldsymbol{E}_{2j}.$$

从而与 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 相应的矩阵就是

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \boldsymbol{D}_1 & & \\ & \ddots & \\ & & \boldsymbol{D}_n \end{pmatrix},$$

其中 $\boldsymbol{D}_j$ 为如下的 2×2 矩阵

$$\boldsymbol{D}_j = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}.$$

这就是具有不同非实特征值矩阵的标准形.

将本节和上节的结果结合起来, 我们得到:

定理 假设 $n \times n$ 矩阵 $\boldsymbol{A}$ 具有不同特征值, 则存在线性映射 $\boldsymbol{T}$ 使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & D_1 & & \\ & & & & \ddots & \\ & & & & & D_l \end{pmatrix},$$

其中 $\boldsymbol{D}_j$ 为如下的 2×2 矩阵

$$\boldsymbol{D}_j = \begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}.$$ ■

5.4 基和子空间

为了讨论具有重特征值的矩阵, 我们需要用到更深入的代数概念. 回忆一下, 一个给定的有限向量组的所有可能线性组合的全体就称为 $\mathbb{R}^n$ 的一个子空间. 准确地说, 给定向量 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k \in \mathbb{R}^n$, 集合

$$\mathcal{S} = \{\alpha_1 \boldsymbol{V}_1 + \cdots + \alpha_k \boldsymbol{V}_k | \alpha_j \in \mathbb{R}\}$$

就是 $\mathbb{R}^n$ 的一个子空间. 此时, 我们称 $\mathcal{S}$ 是由 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k$ **张成的**.

定义 设 $\mathcal{S}$ 为 $\mathbb{R}^n$ 的一个子空间, 向量组 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k$ 称为 $\mathcal{S}$ 的一个**基**, 如果这些 $\boldsymbol{V}_j$ 是线性无关的并且张成 $\mathcal{S}$. ■

注意, 子空间总有基, 这是因为, 如果 $\mathcal{S}$ 是由 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k$ 张成的, 则总可以将这个向量组的某些向量扔掉, 最后得到一个张成 $\mathcal{S}$ 的线性无关的子向量组. 具体地说, 如果这些 $\boldsymbol{V}_j$ 不是线性无关的, 则可以找到某个向量, 比如说 $\boldsymbol{V}_k$, 使得

$$\boldsymbol{V}_k = \beta_1 \boldsymbol{V}_1 + \cdots + \beta_{k-1} \boldsymbol{V}_{k-1}.$$

于是, $\mathcal{S}$ 中的任何向量都可以写成 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_{k-1}$ 的线性组合, 也就是说 $\boldsymbol{V}_k$ 是多余的. 继续这个手法, 最终就可得到一个张成 $\mathcal{S}$ 的线性无关的子向量组.

对于我们的目的而言, 更重要的是:

命题 子空间 $\mathcal{S} \subset \mathbb{R}^n$ 的每一个基都具有相同数目的元素.

证明 首先, 注意到, 具有 k 个线性方程, $k+l$ 个未知变量的系统

$$\begin{aligned} a_{11}x_1+\cdots+a_{1(k+l)}x_{k+l} &= 0 \\ &\vdots \\ a_{k1}x_1+\cdots+a_{k(k+l)}x_{k+l} &= 0 \end{aligned}$$

当 $l>0$ 时, 总有非零解. 事实上, 利用行简化, 先将其中的一个未知变量用其他的未知变量解出来, 然后将这个未知变量消去, 进而得到一个有 $k-1$ 个方程, $k+l-1$ 个未知变量的系统. 最后可用归纳法完成证明 (当 $k=1$ 时, 结论显然成立).

现在假设 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_k$ 为子空间 $\mathcal{S}$ 的一个基, 而 $\boldsymbol{W}_1,\cdots,\boldsymbol{W}_{k+l}$ 也是 $\mathcal{S}$ 的一个基, $l>0$. 于是每个 $\boldsymbol{W}_j$ 都是 $\boldsymbol{V}_i$ 的线性组合, 于是就有常数 a_{ij} 满足:

$$\boldsymbol{W}_j=\sum_{i=1}^{k}a_{ij}\boldsymbol{V}_i,\quad 对\ j=1,\cdots,k+l.$$

根据前面的观察, k 个方程的系统

$$\sum_{j=1}^{k+l}a_{ij}x_j=0,\quad 对\ i=1,\cdots,k$$

有一个非零解 $(c_1,\cdots,c_{k+l})$, 从而就有

$$\sum_{j=1}^{k+l}c_j\boldsymbol{W}_j=\sum_{j=1}^{k+l}c_j\left(\sum_{i=1}^{k}a_{ij}\boldsymbol{V}_i\right)=\sum_{i=1}^{k}\left(\sum_{j=1}^{k+l}a_{ij}c_j\right)\boldsymbol{V}_i=0,$$

由此可得 $\boldsymbol{W}_j$ 为线性相关的, 这一矛盾就完成了证明. ■

作为这一结论的推论, 我们就可以定义子空间的**维数**: 子空间 $\mathcal{S}$ 的维数就是它的任意一个基所含向量的个数. 特别地, $\mathbb{R}^n$ 就是它自己的一个子空间, 它的维数显然就是 n. 进一步, 其他子空间的维数都小于 n, 因为否则的话, $\mathbb{R}^n$ 就会有个数大于 n 的线性无关的向量组, 根据上一个命题, 这不可能. 只由$\mathbf{0}$向量构成的集合也是一个子空间, 我们规定这个子空间的维数为零. 我们用 $\dim\mathcal{S}$ 来记子空间 $\mathcal{S}$ 的维数.

例 $\mathbb{R}^n$ 中一条过原点的直线是 $\mathbb{R}^n$ 的一个一维子空间, 这是因为, 如果任取直线上的一个非零向量 $\boldsymbol{V}\in\mathbb{R}^n$, 则直线上的任一个向量都可以唯一地写成 $t\boldsymbol{V},t\in\mathbb{R}$ 的形式. 显然, 单一向量 $\boldsymbol{V}$ 就构成了这个子空间的一个基. ■

例 在 $\mathbb{R}^3$ 中由

$$x+y+z=0$$

所确定的平面 $\mathcal{P}$ 是 $\mathbb{R}^3$ 的一个二维子空间. 向量 $(1,0,-1)$ 和 $(0,1,-1)$ 都属于 $\mathcal{P}$, 而且它们是线性无关的. 如果 $\boldsymbol{W}\in\mathcal{P}$, 则可将 $\boldsymbol{W}$ 写成

$$\boldsymbol{W}=\begin{pmatrix} x \\ y \\ -x-y \end{pmatrix}=x\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}+y\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix},$$

于是这两个向量张成了 $\mathcal{P}$. ■

与平面情形一样, 我们称一个函数 $\boldsymbol{T}:\mathbb{R}^n\to\mathbb{R}^n$ 为线性的, 如果存在一个 $n\times n$ 矩阵 $\boldsymbol{A}$ 使得 $\boldsymbol{T}(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}$; 此时称 $\boldsymbol{T}$ 为一个**线性映射**或**线性变换**. 利用 5.1 节中得到的矩阵性质, 对任意的 $\alpha,\beta\in\mathbb{R}$ 以及 $\boldsymbol{X},\boldsymbol{Y}\in\mathbb{R}^n$, 我们有,

$$\boldsymbol{T}(\alpha\boldsymbol{X}+\beta\boldsymbol{Y})=\alpha\boldsymbol{T}(\boldsymbol{X})+\beta\boldsymbol{T}(\boldsymbol{Y}).$$

我们称线性映射 $\boldsymbol{T}$ 为可逆的, 如果相应的矩阵 $\boldsymbol{A}$ 可逆.

对于研究线性微分方程系统而言, 最重要的子空间是线性映射的核和值域. 所有被 $\boldsymbol{T}$ 映为 $\boldsymbol{0}$ 的向量构成的集合称为 $\boldsymbol{T}$ 的**核**, 记为 $\mathrm{Ker}\,\boldsymbol{T}$. $\boldsymbol{T}$ 的**值域**(记为 $\mathrm{Range}\,\boldsymbol{T}$) 则由所有这样的向量 $\boldsymbol{W}$ 构成: 存在向量 $\boldsymbol{V}$ 使得 $\boldsymbol{T}\boldsymbol{V}=\boldsymbol{W}$. 显然, 这是微积分中一个熟悉的概念, 这里不同的是, $\boldsymbol{T}$ 的值域总是 $\mathbb{R}^n$ 的一个子空间.

例 考虑线性映射

$$\boldsymbol{T}(\boldsymbol{X})=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\boldsymbol{X}.$$

如果 $\boldsymbol{X}=(x,y,z)$, 则

$$\boldsymbol{T}(\boldsymbol{X})=\begin{pmatrix} y \\ z \\ 0 \end{pmatrix}.$$

于是, $\mathrm{Ker}\,\boldsymbol{T}$ 由所有形如 $(\alpha,0,0)$ 的向量构成, 而 $\mathrm{Range}\,\boldsymbol{T}$ 则由所有形如 $(\beta,\gamma,0)$ 的向量构成, 这里 $\alpha,\beta,\gamma\in\mathbb{R}$. 显然, 这两个集合都是子空间. ■

例 令

$$\boldsymbol{T}(\boldsymbol{X})=\boldsymbol{A}\boldsymbol{X}=\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}\boldsymbol{X}.$$

为了求 $\mathrm{Ker}\,\boldsymbol{T}$, 我们来寻找满足 $\boldsymbol{A}\boldsymbol{X}=\boldsymbol{0}$ 的向量 $\boldsymbol{X}$. 通过行简化, 我们可以将 $\boldsymbol{A}$ 化

成行简化阶梯形矩阵

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

从而 $\boldsymbol{AX} = \mathbf{0}$ 的解 $\boldsymbol{X} = (x, y, z)$ 应满足 $x = z, y = -2z$. 由此可得 $\operatorname{Ker}\boldsymbol{T}$ 中的向量具有形式 $(z, -2z, z)$, 从而 $\operatorname{Ker}\boldsymbol{T}$ 是一维的. 为了求 $\operatorname{Range}\boldsymbol{T}$, 注意到, $\boldsymbol{A}$ 的列向量都在 $\operatorname{Range}\boldsymbol{T}$ 中, 这是因为它们分别是 $(1,0,0), (0,1,0), (0,0,1)$ 的像. 由于

$$-1\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix},$$

这些列向量并不是线性无关的. 不管怎样, $(1,4,7)$ 和 $(2,5,8)$ 是线性无关的, 从而这两个向量构成了 $\operatorname{Range}\boldsymbol{T}$ 的一个基. ■

命题 设 $\boldsymbol{T}: \mathbb{R}^n \to \mathbb{R}^n$ 为一线性映射, 则 $\operatorname{Ker}\boldsymbol{T}$ 和 $\operatorname{Range}\boldsymbol{T}$ 都是 $\mathbb{R}^n$ 的子空间, 而且还有

$$\dim \operatorname{Ker}\boldsymbol{T} + \dim \operatorname{Range}\boldsymbol{T} = n.$$

证明 先假设 $\operatorname{Ker}\boldsymbol{T} = \{\mathbf{0}\}$. 记 $\boldsymbol{E}_1, \cdots, \boldsymbol{E}_n$ 为 $\mathbb{R}^n$ 的标准基. 我们断言 $\boldsymbol{TE}_1, \cdots, \boldsymbol{TE}_n$ 是线性无关的. 若不然, 则存在不全为 0 的 $\alpha_1, \cdots, \alpha_n$, 使得

$$\sum_{j=1}^{n} \alpha_j \boldsymbol{TE}_j = \mathbf{0}.$$

于是就有

$$\boldsymbol{T}\left(\sum_{j=1}^{n} \alpha_j \boldsymbol{E}_j\right) = \mathbf{0},$$

这意味着 $\sum \alpha_j \boldsymbol{E}_j \in \operatorname{Ker}\boldsymbol{T}$, 从而 $\sum \alpha_j \boldsymbol{E}_j = \mathbf{0}$, 于是就有每一个 $\alpha_j = 0$, 这是一个矛盾. 于是, $\boldsymbol{TE}_j$ 为线性无关的. 从而, 任给 $\boldsymbol{V} \in \mathbb{R}^n$, 都存在 $\beta_1, \cdots, \beta_n \in \mathbb{R}$, 使得 $\boldsymbol{V}$ 可以写成

$$\boldsymbol{V} = \sum_{j=1}^{n} \beta_j \boldsymbol{TE}_j,$$

也就是

$$\boldsymbol{V} = \boldsymbol{T}\left(\sum_{j=1}^{n} \beta_j \boldsymbol{E}_j\right),$$

这表明 $\mathrm{Range}\,\boldsymbol{T} = \mathbb{R}^n$. 由此可得, $\mathrm{Ker}\,\boldsymbol{T}$ 和 $\mathrm{Range}\,\boldsymbol{T}$ 都是 $\mathbb{R}^n$ 的子空间, 并且 $\dim\mathrm{Ker}\,\boldsymbol{T} = 0, \dim\mathrm{Range}\,\boldsymbol{T} = n$.

如果 $\mathrm{Ker}\,\boldsymbol{T} \neq \{\boldsymbol{0}\}$, 则可以找到一个非零向量 $\boldsymbol{V}_1 \in \mathrm{Ker}\,\boldsymbol{T}$. 于是, 对所有的 $\alpha \in \mathbb{R}$, $\boldsymbol{T}(\alpha\boldsymbol{V}_1) = \boldsymbol{0}$, 从而所有这样的向量 $\alpha\boldsymbol{V}_1$ 都在 $\mathrm{Ker}\,\boldsymbol{T}$ 中. 如果 $\mathrm{Ker}\,\boldsymbol{T}$ 中还有其他的向量, 任取一个, 把它记为 $\boldsymbol{V}_2$. 由于

$$\boldsymbol{T}(\alpha_1\boldsymbol{V}_1 + \alpha_2\boldsymbol{V}_2) = \alpha_1\boldsymbol{T}\boldsymbol{V}_1 + \alpha_2\boldsymbol{T}\boldsymbol{V}_2 = \boldsymbol{0},$$

$\mathrm{Ker}\,\boldsymbol{T}$ 包含 $\boldsymbol{V}_1, \boldsymbol{V}_2$ 的所有线性组合. 如此继续下去, 我们就可得到一组线性无关的向量, 而且它们张成了 $\mathrm{Ker}\,\boldsymbol{T}$, 这就证明了 $\mathrm{Ker}\,\boldsymbol{T}$ 为一个子空间. 注意, 由于 $\mathbb{R}^n$ 中大于 n 个的向量组是线性相关的, 上面的过程一定会终止. 同样的讨论可说明 $\mathrm{Range}\,\boldsymbol{T}$ 也是一个子空间.

现在假设 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k$ 为 $\mathrm{Ker}\,\boldsymbol{T}$ 的一个基, 其中 $0 < k < n$ (当 $k = n$ 时, 结论显然). 选取向量 $\boldsymbol{W}_{k+1}, \cdots, \boldsymbol{W}_n$ 使得 $\boldsymbol{V}_1, \cdots, \boldsymbol{V}_k, \boldsymbol{W}_{k+1}, \cdots, \boldsymbol{W}_n$ 构成了 $\mathbb{R}^n$ 的一个基. 对每个 j, 记 $\boldsymbol{Z}_j = \boldsymbol{T}\boldsymbol{W}_j$. 则这些向量 $\boldsymbol{Z}_j$ 是线性无关的, 这是因为, 如果有

$$\alpha_{k+1}\boldsymbol{Z}_{k+1} + \cdots + \alpha_n\boldsymbol{Z}_n = 0,$$

则有

$$\boldsymbol{T}(\alpha_{k+1}\boldsymbol{W}_{k+1} + \cdots + \alpha_n\boldsymbol{W}_n) = 0,$$

这将意味着

$$\alpha_{k+1}\boldsymbol{W}_{k+1} + \cdots + \alpha_n\boldsymbol{W}_n \in \mathrm{Ker}\,\boldsymbol{T}.$$

因为我们不能将 $\boldsymbol{W}_j$(以及 $\boldsymbol{W}_j$ 的任一线性组合) 写成 $\boldsymbol{V}_i$ 的线性组合, 所以这是不可能的. 这就证明了 $\mathrm{Ker}\,\boldsymbol{T}$ 和 $\mathrm{Range}\,\boldsymbol{T}$ 的维数之和为 n. ■

需要说明的是, 我们很容易就能找到张成 $\mathrm{Range}\,\boldsymbol{T}$ 的向量组, 只须取与 $\boldsymbol{T}$ 相应的矩阵的列向量组就可以了. 这是因为矩阵的第 i 个列向量就是标准基中的向量 $\boldsymbol{E}_i$ 在 $\boldsymbol{T}$ 下的像. 特别地, 如果这些列向量为线性无关的, 则 $\mathrm{Ker}\,\boldsymbol{T} = \{\boldsymbol{0}\}$, 从而就有, 对每个 $\boldsymbol{V} \in \mathbb{R}^n$, 方程 $\boldsymbol{T}(\boldsymbol{X}) = \boldsymbol{V}$ 有唯一解. 于是, 我们得到:

推论 如果 $\boldsymbol{T}: \mathbb{R}^n \to \mathbb{R}^n$ 为一线性映射, 且 $\dim\mathrm{Ker}\,\boldsymbol{T} = 0$, 则 $\boldsymbol{T}$ 可逆. ■

5.5 重特征值

在本节中, 我们来描述当矩阵有重特征值时的标准形. 为了不在发展一般理论时花费过多的时间, 我们将只对有重特征值的 3×3 和 4×4 矩阵作详细讨论. 对于更一般的情形则留作习题. 我们将在下节说明这种省略是合理的, 因为在那里我们将说明"典型"的矩阵都具有不同特征值, 从而可以像前面的章节中同样处

理.(如果你在逛街时随机碰到一个矩阵, 则它就很有可能具有不同特征值.) 关于具有重特征值矩阵最一般的结论是:

命题　设 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵, 则存在坐标变换 $\boldsymbol{T}$ 使得

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} \boldsymbol{B}_1 & & \\ & \ddots & \\ & & \boldsymbol{B}_k \end{pmatrix},$$

其中每个 $\boldsymbol{B}_j$ 都是一个方阵 (而其他所有的元素都是零), 它的形式为以下两者之一:

$$\text{(i)} \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \qquad \text{(ii)} \begin{pmatrix} \boldsymbol{C}_2 & \boldsymbol{I}_2 & & & \\ & \boldsymbol{C}_2 & \boldsymbol{I}_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \boldsymbol{I}_2 \\ & & & & \boldsymbol{C}_2 \end{pmatrix},$$

其中

$$\boldsymbol{C}_2 = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \quad \boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

并且这里的 $\alpha, \beta, \lambda \in \mathbb{R}$, $\beta \neq 0$. 此外, 特别情形 $\boldsymbol{B}_j = (\lambda)$ 或

$$\boldsymbol{B}_j = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

当然也是容许的. ■

先考虑 $\mathbb{R}^3$ 的情形. 如果 $\boldsymbol{A}$ 在 $\mathbb{R}^3$ 中具有重特征值, 则所有的特征值都是实的. 此时有两种情形, 要么有两个不同的特征值, 其中一个是重的, 要么所有的特征值都相同. 前一种情形可以利用在第 3 章中所描述的类似过程进行处理, 因而我们只关心 $\boldsymbol{A}$ 仅有一个三重特征值 λ 的情形.

命题　假设 $\boldsymbol{A}$ 为一 3×3 矩阵, 具有唯一的特征值 λ. 则可以找到坐标变换使得 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 为下面的三种形式之一:

$$\text{(i)} \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{(ii)} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{(iii)} \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}.$$

证明　记 K 为 $\boldsymbol{A} - \lambda\boldsymbol{I}$ 的核. K 中任一非零向量都是 $\boldsymbol{A}$ 的特征向量. 根据 K 的维数分别为 $1, 2$ 或 3, 我们有对应的三种情形.

当 K 的维数为 3 时, 则对所有的 $\boldsymbol{V} \in \mathbb{R}^3$, $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}=\boldsymbol{0}$, 从而 $\boldsymbol{A}=\lambda\boldsymbol{I}$, 这就得到情形 (i) 中的矩阵.

假设 K 的维数为 2. 记 R 为 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 的值域. 由上节可知 $\dim K+\dim R=3$, 从而 R 的维数为 1. 我们断言 $R\subset K$. 若不然, 任取非零向量 $\boldsymbol{V}\in R$, 由于 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}\in R$ 且 R 为一维的, 因而存在 $\mu\in\mathbb{R}, \mu\neq 0$, 使得 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}=\mu\boldsymbol{V}$. 于是 $\boldsymbol{AV}=(\lambda+\mu)\boldsymbol{V}$, 这样我们就找到了一个新的特征值 $\lambda+\mu$, 这与假设矛盾, 从而我们必有 $R\subset K$.

现在, 取非零的 $\boldsymbol{V}_1\in R$, 由于 $\boldsymbol{V}_1\in K$, $\boldsymbol{V}_1$ 为一特征向量, 从而 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}_1=\boldsymbol{0}$, 又由于 $\boldsymbol{V}_1$ 也在 R 中, 我们可找到 $\boldsymbol{V}_2\in\mathbb{R}^3-K$ 使得 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}_2=\boldsymbol{V}_1$. 因为 K 为二维的, 我们可以再取第二个向量 $\boldsymbol{V}_3\in K$ 使得 $\boldsymbol{V}_1$ 和 $\boldsymbol{V}_3$ 为线性无关的. 注意, $\boldsymbol{V}_3$ 也是一个特征向量. 现在, 如果做坐标变换使得 $\boldsymbol{TE}_j=\boldsymbol{V}_j, j=1,2,3$, 则易见 $T^{-1}\boldsymbol{A}T$ 为情形 (ii) 中的形式.

最后, 假设 K 的维数等于 1, 从而 R 的维数等于 2. 我们断言, 此时 $K\subset R$. 若不然, 则有 $(\boldsymbol{A}-\lambda\boldsymbol{I})R=R$, 从而 $\boldsymbol{A}-\lambda\boldsymbol{I}$ 限制在 R 上可逆. 于是, 如果 $\boldsymbol{V}\in R$, 则存在唯一的 $\boldsymbol{W}\in R$ 满足 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{W}=\boldsymbol{V}$. 特别地, 我们有

$$\boldsymbol{AV}=\boldsymbol{A}(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{W}=(\boldsymbol{A}^2-\lambda\boldsymbol{A})\boldsymbol{W}=(\boldsymbol{A}-\lambda\boldsymbol{I})(\boldsymbol{AW}).$$

这表明如果 $\boldsymbol{V}\in R$, 则 $\boldsymbol{AV}$ 也在 R 中. 从而 $\boldsymbol{A}$ 也保持子空间 R 不变. 于是立刻就有, $\boldsymbol{A}$ 在 R 中必有特征向量, 而这将意味着 $K\subset R$, 矛盾.

接下来, 我们断言 $(\boldsymbol{A}-\lambda\boldsymbol{I})R=K$. 为此, 先由 $K\subset R$ 可得 $(\boldsymbol{A}-\lambda\boldsymbol{I})R$ 为一维的. 如果 $(\boldsymbol{A}-\lambda\boldsymbol{I})R\neq K$, 则存在非零向量 $\boldsymbol{V}\notin K$, 使得 $(\boldsymbol{A}-\lambda\boldsymbol{I})R=\{t\boldsymbol{V}|t\in\mathbb{R}\}$, 于是就存在 $t\in\mathbb{R}, t\neq 0$ 满足 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}=t\boldsymbol{V}$, 即 $\boldsymbol{AV}=(t+\lambda)\boldsymbol{V}$, 这样就得到了另一个新特征值. 因而必有 $(\boldsymbol{A}-\lambda\boldsymbol{I})R=K$.

现在令 $\boldsymbol{V}_1\in K$ 为 $\boldsymbol{A}$ 的一个特征向量, 由上面的结论可知, 存在 $\boldsymbol{V}_2\in R$ 使得 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}_2=\boldsymbol{V}_1$. 又因为 $\boldsymbol{V}_2\in R$, 故存在 $\boldsymbol{V}_3$ 使得 $(\boldsymbol{A}-\lambda\boldsymbol{I})\boldsymbol{V}_3=\boldsymbol{V}_2$, 可见 $(\boldsymbol{A}-\lambda\boldsymbol{I})^2\boldsymbol{V}_3=\boldsymbol{V}_1$. 易见, $\boldsymbol{V}_j$ 为线性无关的, 而且由 $\boldsymbol{TE}_j=\boldsymbol{V}_j$ 定义的线性映射可将 $\boldsymbol{A}$ 化为 (iii) 中的标准形. 证毕. ■

例 设

$$\boldsymbol{A}=\begin{pmatrix} 2 & 0 & -1 \\ 0 & 2 & 1 \\ -1 & -1 & 2 \end{pmatrix}.$$

按第一行展开可得

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(2-\lambda)[(2-\lambda)^2+1]-(2-\lambda)=(2-\lambda)^3,$$

于是 2 是唯一的特征值. 解 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{V}=\boldsymbol{0}$ 可得只有一个线性无关的特征向量 $\boldsymbol{V}_1=(1,-1,0)$, 这属于命题中的情形 (iii). 经过计算, 我们有

$$(\boldsymbol{A}-2\boldsymbol{I})^2=\begin{pmatrix}1&1&0\\-1&-1&0\\0&0&0\end{pmatrix},$$

因而向量 $\boldsymbol{V}_3=(1,0,0)$ 满足 $(\boldsymbol{A}-2\boldsymbol{I})^2\boldsymbol{V}_3=\boldsymbol{V}_1$. 同时, 我们还有

$$(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{V}_3=\boldsymbol{V}_2=(0,0,-1).$$

如前, 令 $T\boldsymbol{E}_j=\boldsymbol{V}_j, j=1,2,3$, 从而就有

$$\boldsymbol{T}=\begin{pmatrix}1&0&1\\-1&0&0\\0&-1&0\end{pmatrix},$$

而 $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ 具有标准形

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}=\begin{pmatrix}2&1&0\\0&2&1\\0&0&2\end{pmatrix}.$$ ■

例 设

$$\boldsymbol{A}=\begin{pmatrix}1&1&0\\-1&3&0\\-1&1&2\end{pmatrix}.$$

还是按第一行展开, 可得

$$\det(\boldsymbol{A}-\lambda\boldsymbol{I})=(1-\lambda)[(3-\lambda)(2-\lambda)]+(2-\lambda)=(2-\lambda)^3,$$

于是 2 也是唯一的特征值. 然而, 此时, 我们有

$$\boldsymbol{A}-2\boldsymbol{I}=\begin{pmatrix}-1&1&0\\-1&1&0\\-1&1&0\end{pmatrix},$$

于是我们有两个线性无关的特征向量 (x,y,z), 其中 $x=y$, 而 z 是任意的. 注意, 此时 $(\boldsymbol{A}-2\boldsymbol{I})^2$ 为零矩阵, 因而我们可以任取一个非特征向量的向量作为 $\boldsymbol{V}_2$, 比如, 可取 $\boldsymbol{V}_2=(1,0,0)$. 于是 $(\boldsymbol{A}-2\boldsymbol{I})\boldsymbol{V}_2=\boldsymbol{V}_1=(-1,-1,-1)$ 为一特征向量, 另一个

线性无关的特征向量可取 $\boldsymbol{V}_3=(0,0,1)$. 按通常的方式定义 $\boldsymbol{TE}_j=\boldsymbol{V}_j$ 后, 可得标准形

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}2&1&0\\0&2&0\\0&0&2\end{pmatrix}.$$

■

现在我们来看 4×4 的情形. 当有实特征值时, 情况与 3×3 的情形是类似的 (但代数上更复杂), 因而将它留作习题. 于是下面我们假设 $\boldsymbol{A}$ 具有重的复特征值 $\alpha\pm \mathrm{i}\beta$, $\beta\neq 0$.

只有两种情形: 属于 $\alpha+\mathrm{i}\beta$ 的线性无关特征向量要么有一对, 要么只有一个. 在前一种情形, 令 $\boldsymbol{V}_1,\boldsymbol{V}_2$ 为线性无关的特征向量, 则 $\overline{\boldsymbol{V}}_1,\overline{\boldsymbol{V}}_2$ 就是属于 $\alpha-\mathrm{i}\beta$ 的两个线性无关的特征向量. 像前面一样, 如下选取实向量

$$\begin{aligned}\boldsymbol{W}_1&=(\boldsymbol{V}_1+\overline{\boldsymbol{V}}_1)/2\\ \boldsymbol{W}_2&=-\mathrm{i}(\boldsymbol{V}_1-\overline{\boldsymbol{V}}_1)/2\\ \boldsymbol{W}_3&=(\boldsymbol{V}_2+\overline{\boldsymbol{V}}_2)/2\\ \boldsymbol{W}_4&=-\mathrm{i}(\boldsymbol{V}_2-\overline{\boldsymbol{V}}_2)/2.\end{aligned}$$

如果令 $\boldsymbol{TE}_j=\boldsymbol{W}_j$, 则坐标变换 $\boldsymbol{T}$ 可将 $\boldsymbol{A}$ 化为标准形

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}\alpha&\beta&0&0\\-\beta&\alpha&0&0\\0&0&\alpha&\beta\\0&0&-\beta&\alpha\end{pmatrix}.$$

如果只能找到 $\alpha+\mathrm{i}\beta$ 的一个线性无关的特征向量 $\boldsymbol{V}_1$, 像在重实特征值情形一样, 我们来求解方程组系统 $(\boldsymbol{A}-(\alpha+\mathrm{i}\beta)\boldsymbol{I})\boldsymbol{X}=\boldsymbol{V}_1$. 从上一个命题的证明中可以看出, 我们总可以找到这个方程组的一个非零解 $\boldsymbol{V}_2$. 然后像上面一样选取 $\boldsymbol{W}_j$, 再令 $\boldsymbol{TE}_j=\boldsymbol{W}_j$, 则 $\boldsymbol{T}$ 就将 $\boldsymbol{A}$ 化为标准形

$$\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}\alpha&\beta&1&0\\-\beta&\alpha&0&1\\0&0&\alpha&\beta\\0&0&-\beta&\alpha\end{pmatrix}.$$

作为一个例子, 我们做下面的计算:

$$\begin{aligned}(T^{-1}AT)E_3 &= T^{-1}AW_3\\ &= T^{-1}A(V_2+\overline{V}_2)/2\\ &= T^{-1}\big((V_1+(\alpha+\mathrm{i}\beta)V_2)/2+(\overline{V}_1+(\alpha-\mathrm{i}\beta)\overline{V}_2)/2\big)\\ &= T^{-1}\big((V_1+\overline{V}_1)/2+\alpha(V_2+\overline{V}_2)/2+\mathrm{i}\beta(V_2-\overline{V}_2)/2\big)\\ &= E_1+\alpha E_3-\beta E_4.\end{aligned}$$

例 令

$$A=\begin{pmatrix}1 & -1 & 0 & 1\\ 2 & -1 & 1 & 0\\ 0 & 0 & -1 & 2\\ 0 & 0 & -1 & 1\end{pmatrix}.$$

经过一些计算后, 可得其特征方程为

$$(\lambda^2+1)^2=0.$$

因而 A 有重特征值 $\pm\mathrm{i}$, 且每一个都是两重的.

求解系统 $(A-\mathrm{i}I)X=\mathbf{0}$ 可得 i 的一个线性无关的复特征向量 $V_1=(1,1-i,0,0)$, 从而 $\overline{V}_1$ 是属于 $-\mathrm{i}$ 的一个特征向量.

然后求解系统 $(A-\mathrm{i}I)X=V_1$ 可得 $V_2=(0,0,1-\mathrm{i},1)$, 于是 $\overline{V}_2$ 为系统 $(A-\mathrm{i}I)X=\overline{V}_1$ 的一个解. 最后, 我们令

$$\begin{aligned}W_1 &= (V_1+\overline{V}_1)/2=\mathrm{Re}\,V_1\\ W_2 &= -\mathrm{i}(V_1-\overline{V}_1)/2=\mathrm{Im}\,V_1\\ W_3 &= (V_2+\overline{V}_2)/2=\mathrm{Re}\,V_2\\ W_4 &= -\mathrm{i}(V_2-\overline{V}_2)/2=\mathrm{Im}\,V_2,\end{aligned}$$

以及 $TE_j=W_j, j=1,2,3,4$. 我们有

$$T=\begin{pmatrix}1 & 0 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 1 & 0\end{pmatrix},\quad T^{-1}=\begin{pmatrix}1 & 0 & 0 & 0\\ 1 & -1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & -1 & 1\end{pmatrix},$$

从而得到标准形

$$T^{-1}AT=\begin{pmatrix}0&1&1&0\\-1&0&0&1\\0&0&0&1\\0&0&-1&0\end{pmatrix}.$$

■

例 令

$$A=\begin{pmatrix}2&0&1&0\\0&2&0&1\\0&0&2&0\\0&-1&0&2\end{pmatrix}.$$

A 的特征方程为

$$(2-\lambda)^2((2-\lambda)^2+1)=0,$$

于是特征值为 $2\pm\mathrm{i}$ 和 2(其重数为 2).

解方程组 $(A-(2+\mathrm{i})I)X=0$ 可得 $2+\mathrm{i}$ 的一个特征向量为 $V=(0,-\mathrm{i},0,1)$. 令 $W_1=(0,0,0,1), W_2=(0,-1,0,0)$, 它们分别是 V 的实部和虚部.

解方程组 $(A-2I)X=0$ 只可得到属于 2 的一个特征向量, 即 $W_3=(1,0,0,0)$. 这样, 我们再求解 $(A-2I)X=W_3$ 可找到 $W_4=(0,0,1,0)$. 像通常一样, 令 $TE_j=W_j$, 则容易验证, T 将 A 化为标准形

$$T^{-1}AT=\begin{pmatrix}2&1&0&0\\-1&2&0&0\\0&0&2&1\\0&0&0&2\end{pmatrix}.$$

■

5.6 通 有 性

我们已经多次提到过, "大多数" 矩阵具有不同特征值, 本节的目标就是使这句话严格化.

先复习 $\mathbb{R}^n$ 中的一些拓扑概念. 一个集合 $\mathcal{U}\subset\mathbb{R}^n$ 称为**开的**, 如果对任意的 $X\in\mathcal{U}$, 总有 X 的一个开球包含在 $\mathcal{U}$ 中, 即, 存在某个 $a>0$(依赖于 X), 使得 X 的半径为 a 的开球

$$\{Y\in\mathbb{R}^n \,\big|\, |Y-X|<a\}$$

包含在 $\mathcal{U}$ 中. 用几何语言, 可以说成: 如果 X 属于一个开集 $\mathcal{U}$, 则充分靠近 X 的点也在 $\mathcal{U}$ 中.

$\mathbb{R}^n$ 中的另外一种子集是所谓的**稠密集**：$\mathcal{U} \subset \mathbb{R}^n$ 称为稠密的，如果 $\mathbb{R}^n$ 中的每一个点都可以被 $\mathcal{U}$ 中的点任意接近. 精确地说，如果 $\boldsymbol{X} \in \mathbb{R}^n$，则对每一个 $\epsilon > 0$，都存在某个 $\boldsymbol{Y} \in \mathcal{U}$，使得 $|\boldsymbol{X} - \boldsymbol{Y}| < \epsilon$. $\mathcal{U}$ 在 $\mathbb{R}^n$ 中稠密的一个等价描述是：对于每一个非空开集 $\mathcal{V} \subset \mathbb{R}^n$，$\mathcal{V} \cap \mathcal{U}$ 都是非空的. 例如，全体有理数构成了 $\mathbb{R}$ 的一个稠密子集，全体无理数也是如此. 类似地，

$$\{(x, y) \in \mathbb{R}^2 | x, y\text{都是有理数}\}$$

是平面的一个稠密子集.

$\mathbb{R}^n$ 中一类有趣的子集是所谓的既开又稠密的集合. 这样的一个集合 $\mathcal{U}$ 可以由如下的性质来刻画：$\mathcal{U}$ 的余集中的点可以被 $\mathcal{U}$ 中的点任意逼近 (因为 $\mathcal{U}$ 稠密)，但是 $\mathcal{U}$ 没有点可以被 $\mathcal{U}$ 的余集中的点任意逼近 (因为 $\mathcal{U}$ 是开集).

$\mathbb{R}^2$ 中的一个简单的开稠集如下：

$$\mathcal{V} = \{(x, y) \in \mathbb{R}^2 | xy \neq 1\}.$$

显然，这是 $\mathbb{R}^2$ 中由 $xy = 1$ 所定义的双曲线的余集. 假设 $(x_0, y_0) \in \mathcal{V}$. 于是，$x_0 y_0 \neq 1$，从而当 $|x - x_0|, |y - y_0|$ 都充分小时，也有 $xy \neq 1$；这就证明了 $\mathcal{V}$ 为开的. 任给 $(x_0, y_0) \in \mathbb{R}^2$，我们可以找到任意靠近 (x_0, y_0) 的 (x, y) 使得 $xy \neq 1$；这就证明了 $\mathcal{V}$ 是稠密的.

下面的命题说明，开稠集是一种很 "胖" 的集合.

命题　设 $\mathcal{V}_1, \cdots, \mathcal{V}_m$ 都是 $\mathbb{R}^n$ 中的开稠集，则

$$\mathcal{V} = \mathcal{V}_1 \cap \cdots \cap \mathcal{V}_m$$

也是开稠的.

证明　易见，有限个开集的交仍然是开的，因而 $\mathcal{V}$ 是开的. 为了证明 $\mathcal{V}$ 的稠密性，任取一个非空开集 $\mathcal{U} \subset \mathbb{R}^n$，因为 $\mathcal{V}_1$ 稠密，$\mathcal{U} \cap \mathcal{V}_1$ 非空. 又因为 $\mathcal{U}, \mathcal{V}_1$ 都是开的，所以 $\mathcal{U} \cap \mathcal{V}_1$ 也是开的. 由于 $\mathcal{U} \cap \mathcal{V}_1$ 是非空开集，再加上 $\mathcal{V}_2$ 的稠密性，可得 $(\mathcal{U} \cap \mathcal{V}_1) \cap \mathcal{V}_2$ 也是非空的. 而 $\mathcal{V}_2$ 为开意味着 $\mathcal{U} \cap \mathcal{V}_1 \cap \mathcal{V}_2$ 也是开的. 于是，$(\mathcal{U} \cap \mathcal{V}_1 \cap \mathcal{V}_2) \cap \mathcal{V}_3$ 是非空的，依次这么下去，最后可知，$\mathcal{U} \cap \mathcal{V}$ 是非空的，这就证明了 $\mathcal{V}$ 在 $\mathbb{R}^n$ 中稠密. ■

于是我们可以这样认为，$\mathbb{R}^n$ 中的子集如果含有一个开稠集，则它是大的. 为了使 "大多数" 矩阵这一说法严格化，我们需要在所有矩阵的集合中引入开稠集的概念.

记 $L(\mathbb{R}^n)$ 为 $n \times n$ 矩阵的集合，这也就是 $\mathbb{R}^n$ 上所有线性映射的集合. 为了讨论 $L(\mathbb{R}^n)$ 中的开稠集，我们需要有一个度量 $L(\mathbb{R}^n)$ 中两个矩阵远近的概念. 这件事很容易办到，我们只须将一个矩阵的所有元素写成 (按一种指定的顺序) 一个长向量就可以了，由此，可将 $L(\mathbb{R}^n)$ 看作是 $\mathbb{R}^{n^2}$.

定理 $L(\mathbb{R}^n)$ 中由具有 n 个不同特征值的矩阵所构成的集合 $\mathcal{M}$ 在 $L(\mathbb{R}^n)$ 中是开稠的.

证明 我们首先证明 $\mathcal{M}$ 的稠密性. 设 $\boldsymbol{A} \in L(\mathbb{R}^n)$, 假定 $\boldsymbol{A}$ 有一些重特征值. 上节的命题说明, 我们可以找到矩阵 $\boldsymbol{T}$ 使得 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 具有如下的两种形式之一：要么标准形是分块对角矩阵, 且每个矩阵块具有形式

$$\text{(i)} \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix} \quad \text{或} \quad \text{(ii)} \begin{pmatrix} \boldsymbol{C}_2 & \boldsymbol{I}_2 & & & \\ & \boldsymbol{C}_2 & \boldsymbol{I}_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & \boldsymbol{I}_2 \\ & & & & \boldsymbol{C}_2 \end{pmatrix},$$

其中 $\alpha, \beta, \lambda \in \mathbb{R}$, $\beta \neq 0$, 且

$$\boldsymbol{C}_2 = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}, \quad \boldsymbol{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

要么是一对孤立对角块 (λ) 或 $\boldsymbol{C}_2$. 每种情形都可以处理如下.

取一些不同的值 λ_j, 使得 $|\lambda - \lambda_j|$ 像事先要求的那样小, 然后将 (i) 中的块替换成

$$\begin{pmatrix} \lambda_1 & 1 & & & \\ & \lambda_2 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda_j \end{pmatrix}.$$

这个新块具有不同的特征值. 对于 (ii) 中的块, 我们同样类似地将 2×2 的块

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

中 α 替换成一些不同的 α_i. 新矩阵就具有不同的特征值 $\alpha_i \pm \mathrm{i}\beta$. 如此, 我们就找到了一个新矩阵 $\boldsymbol{B}$ 充分接近 $\boldsymbol{T}^{-1}\boldsymbol{AT}$, 并且有不同的特征值. 于是矩阵 $\boldsymbol{TBT}^{-1}$ 也就具有不同的特征值, 并且任意接近 $\boldsymbol{A}$. 事实上, 当 $\boldsymbol{T}$ 是一个固定的可逆矩阵时, 由 $F(\boldsymbol{M}) = \boldsymbol{TMT}^{-1}$ 定义的函数 $F: L(\mathbb{R}^n) \to L(\mathbb{R}^n)$ 是 $L(\mathbb{R}^n)$ 上的一个连续函数, 因而它将接近 $\boldsymbol{T}^{-1}\boldsymbol{AT}$ 的矩阵带到接近 $\boldsymbol{A}$ 的新矩阵. 这就证明了 $\mathcal{M}$ 的稠密性.

为了证明 $\mathcal{M}$ 的开性, 我们来考虑矩阵 $\boldsymbol{A} \in L(\mathbb{R}^n)$ 的特征多项式. 如果将 $\boldsymbol{A}$ 的元素作微小的改变, 则特征多项式的系数变化也很小, 从而这个多项式在 $\mathbb{C}$ 中的根的移动也很小. 因而, 如果从一个具有不同特征值的矩阵出发, 它附近的矩阵也有这个性质. 这就证明了 $\mathcal{M}$ 的开性. ■

矩阵的一个性质 $\mathcal{P}$ 称为一个**通有性质**, 如果具有性质 $\mathcal{P}$ 的矩阵的集合包含 $L(\mathbb{R}^n)$ 中的一个开稠集. 从而, 一个性质是通有的, 如果矩阵的某一开稠集都具有这个性质 (其他矩阵也可能具有). 直观地说, 通有性质就是指 "几乎所有" 矩阵都具有的性质. 从而, 具有不同特征值是 $n \times n$ 矩阵的一个通有性质.

习　题

1. 证明一个 3×3 矩阵的行列式可以通过按任一行 (列) 展开来计算.
2. 找出下列矩阵的特征值和特征向量:

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$　(b) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 0 \\ 3 & 0 & 0 \end{pmatrix}$　(c) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 0 & 2 \\ 0 & 2 & 0 \\ -2 & 0 & 0 \end{pmatrix}$　(e) $\begin{pmatrix} 3 & 0 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & -2 & -1 & -4 \\ -1 & 0 & 0 & 3 \end{pmatrix}$

3. 在 abc 空间中描述矩阵

$$\begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{pmatrix}$$

分别具有实特征值、复特征值以及重特征值的区域.

4. 在 abc 空间站中描述矩阵

$$\begin{pmatrix} a & 0 & 0 & a \\ 0 & a & b & 0 \\ 0 & c & a & 0 \\ a & 0 & 0 & a \end{pmatrix}$$

分别具有实特征值、复特征值以及重特征值的区域.

5. 将下列矩阵化为标准形

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ (e) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$ (h) $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

6. 假设一个 5×5 矩阵的特征值为 $2, 1 \pm \mathrm{i}$, 列出这种类型矩阵的所有可能标准形.
7. 设 $\boldsymbol{L}$ 是交换一个矩阵的第 i, j 两行的初等矩阵, 即, 除了 $l_{ii} = l_{jj} = 0, l_{ij} = l_{ji} = 1$, $\boldsymbol{L}$ 沿对角线取值为 1. 证明 $\det \boldsymbol{L} = -1$.
8. 当 $\boldsymbol{T}$ 为下列矩阵时, 找出 $\mathrm{Ker}\,\boldsymbol{T}$ 和 $\mathrm{Range}\,\boldsymbol{T}$ 的一个基.

(a) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ (b) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ (c) $\begin{pmatrix} 1 & 9 & 6 \\ 1 & 4 & 1 \\ 2 & 7 & 1 \end{pmatrix}$

9. 假设 4×4 矩阵 $\boldsymbol{A}$ 只有一个实特征值 λ 并且只有一个线性无关的特征向量. 证明可以将 $\boldsymbol{A}$ 化为如下的标准形:

$$\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}.$$

10. 假设 4×4 矩阵 $\boldsymbol{A}$ 只有一个实特征值 λ 并且有两个线性无关的特征向量. 描述 $\boldsymbol{A}$ 所有可能的标准形, 并证明可以将 $\boldsymbol{A}$ 化成这些标准形中的某一个. 并且明确地给出在哪些条件下, $\boldsymbol{A}$ 可以化成某个特定的标准形.
11. 证明: 如果 $\boldsymbol{A}$ 和 (或)$\boldsymbol{B}$ 为不可逆矩阵, 则 $\boldsymbol{AB}$ 也是不可逆的.
12. 假设 $\mathcal{S}$ 为 $\mathbb{R}^n$ 中一个满足下列条件的子集:
    (a) 如果 $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{S}$, 则 $\boldsymbol{X} + \boldsymbol{Y} \in \mathcal{S}$;
    (b) 如果 $\boldsymbol{X} \in \mathcal{S}, \alpha \in \mathbb{R}$, 则 $\alpha \boldsymbol{X} \in \mathcal{S}$.
    证明 $\mathcal{S}$ 可以写成有限个向量的所有可能线性组合的全体.
13. $\mathbb{R}^n$ 中的下列哪些子集是开的? 哪些子集是稠密的? 请给出简要解释.

(a) $\mathcal{U}_1 = \{(x,y)|y>0\}$; (b) $\mathcal{U}_2 = \{(x,y)|x^2+y^2 \neq 1\}$;

(c) $\mathcal{U}_3 = \{(x,y)|x$为无理数$\}$; (d) $\mathcal{U}_4 = \{(x,y)|x,y$都不是整数$\}$;

(e) $\mathcal{U}_5$ 为集合 C_1 的余集, 其中 C_1 是闭的但不稠密;

(f) $\mathcal{U}_6$ 为集合 C_2 的余集, 其中 C_2 恰好含有 60 亿个不同的点.

14. 下列每条性质都定义了实 $n \times n$ 矩阵的一个子集. 哪些子集在 $L(\mathbb{R}^n)$ 中是开的? 哪些子集在 $L(\mathbb{R}^n)$ 中是稠密的? 请给出简要解释.

(a) $\det \boldsymbol{A} \neq 0$. (b)Trace $\boldsymbol{A}$ 为有理数.

(c) $\boldsymbol{A}$ 的元素都不是整数. (d)$3 \leqslant \det \boldsymbol{A} < 4$.

(e) 每个特征值 λ 都满足 $-1 < |\lambda| < 1$. (f)$\boldsymbol{A}$ 没有实特征值.

(g) $\boldsymbol{A}$ 的每个实特征值都是单重的.

15. $\mathbb{R}^n$ 上线性映射的下面哪些性质是通有的?

(a) 每个特征值 λ 都满足 $|\lambda| \neq 1$. (b) $n=2$ 且有非实特征值.

(c) $n=3$ 且有非实特征值. (d) $\boldsymbol{X}' = \boldsymbol{AX}$ 没有周期解 (零解除外).

(e) 有 n 个不同的特征值并且虚部也不相同.

(f) 对所有的 $\boldsymbol{X} \neq \boldsymbol{0}$ 都有 $\boldsymbol{AX} \neq \boldsymbol{X}$且$\boldsymbol{AX} \neq -\boldsymbol{X}$.

第 6 章　高维线性系统

在线性代数领域稍作停留后, 现在该回到微分方程了, 特别地, 要回到求解具有常系数的高维线性系统的任务中来. 和线性代数那章一样, 我们要讨论大量的不同情形.

6.1　不同特征值

首先考虑线性系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 其中 $n \times n$ 矩阵 $\boldsymbol{A}$ 具有 n 个不同的实特征值 $\lambda_1, \cdots, \lambda_n$. 根据第 5 章中的结论, 存在坐标变换 $\boldsymbol{T}$ 使得新系统 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 具有如下相当简单的形式:

$$\begin{aligned} y_1' &= \lambda_1 y_1 \\ &\vdots \\ y_n' &= \lambda_n y_n, \end{aligned}$$

其中的线性映射 $\boldsymbol{T}$ 将标准基向量 $\boldsymbol{E}_j$ 映到属于 λ_j 的特征向量 $\boldsymbol{V}_j$. 显然, 下面形式的函数

$$\boldsymbol{Y}(t) = \begin{pmatrix} c_1 \mathrm{e}^{\lambda_1 t} \\ \vdots \\ c_n \mathrm{e}^{\lambda_n t} \end{pmatrix}$$

是 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 满足初值条件 $\boldsymbol{Y}(0) = (c_1, \cdots, c_n)$ 的一个解. 与第 3 章一样, 它是这样的唯一解, 这是因为, 如果

$$\boldsymbol{W}(t) = \begin{pmatrix} w_1(t) \\ \vdots \\ w_n(t) \end{pmatrix}$$

为另一个解, 则对每个表达式 $w_j(t)\exp(-\lambda_j t)$ 求导可得

$$\frac{\mathrm{d}}{\mathrm{d}t} w_j(t)\mathrm{e}^{-\lambda_j t} = (w_j' - \lambda_j w_j)\mathrm{e}^{-\lambda_j t} = 0.$$

于是对每个 j, 都有 $w_j(t) = c_j \exp(\lambda_j t)$. 因而, 解 $\boldsymbol{Y}(t)$ 的全体就是 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的通解.

由此可得 $\boldsymbol{X}(t)=\boldsymbol{T}\boldsymbol{Y}(t)$ 就是 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的通解, 而且这个通解可以写成如下形式

$$\boldsymbol{X}(t)=\sum_{j=1}^{n}c_j\mathrm{e}^{\lambda_j t}\boldsymbol{V}_j.$$

现在假设 $\boldsymbol{A}$ 的特征值中 $\lambda_1,\cdots,\lambda_k$ 为负的, 而 $\lambda_{k+1},\cdots,\lambda_n$ 为正的. 因为没有零特征值, 故此时系统是双曲的. 首先, 所有从 $\boldsymbol{V}_1,\cdots,\boldsymbol{V}_k$ 张成的子空间出发的解都将永远逗留在这个子空间中, 这是因为此时 $c_{k+1}=\cdots=c_n=0$. 其次, 每个这样的解在 $t\to\infty$ 时都趋于原点. 与平面系统中引入的术语类似, 我们称这个子空间为**稳定子空间**. 类似地, $\boldsymbol{V}_{k+1},\cdots,\boldsymbol{V}_n$ 张成的子空间中的解将远离原点. 这个子空间称为**不稳定子空间**. 所有其他的解在时间向后时都趋于稳定子空间, 而在时间向前时都趋于不稳定子空间. 因而这个系统对应于高维的**鞍点**.

例　考虑

$$\boldsymbol{X}'=\begin{pmatrix}1&2&-1\\0&3&-2\\0&2&-2\end{pmatrix}\boldsymbol{X}.$$

在 5.2 节中, 我们已经看到这个矩阵的特征值为 $2,1,-1$, 属于它们的特征向量分别为 $(3,2,1),(1,0,0),(0,1,2)$. 于是矩阵

$$\boldsymbol{T}=\begin{pmatrix}3&1&0\\2&0&1\\1&0&2\end{pmatrix}$$

就将 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 化为

$$\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}=\begin{pmatrix}2&0&0\\0&1&0\\0&0&-1\end{pmatrix}\boldsymbol{Y},$$

我们可以立刻得到它的解, 用 $\boldsymbol{T}$ 乘以这个解就可以得到 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 的通解为

$$\boldsymbol{X}(t)=c_1\mathrm{e}^{2t}\begin{pmatrix}3\\2\\1\end{pmatrix}+c_2\mathrm{e}^{t}\begin{pmatrix}1\\0\\0\end{pmatrix}+c_3\mathrm{e}^{-t}\begin{pmatrix}0\\1\\2\end{pmatrix}.$$

过原点以及 $(0,1,2)$ 的直线就是稳定线, 而由 $(3,2,1),(1,0,0)$ 张成的平面是不稳定平面. 该系统以及系统 $\boldsymbol{Y}'=(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的解如图 6.1 所示. ■

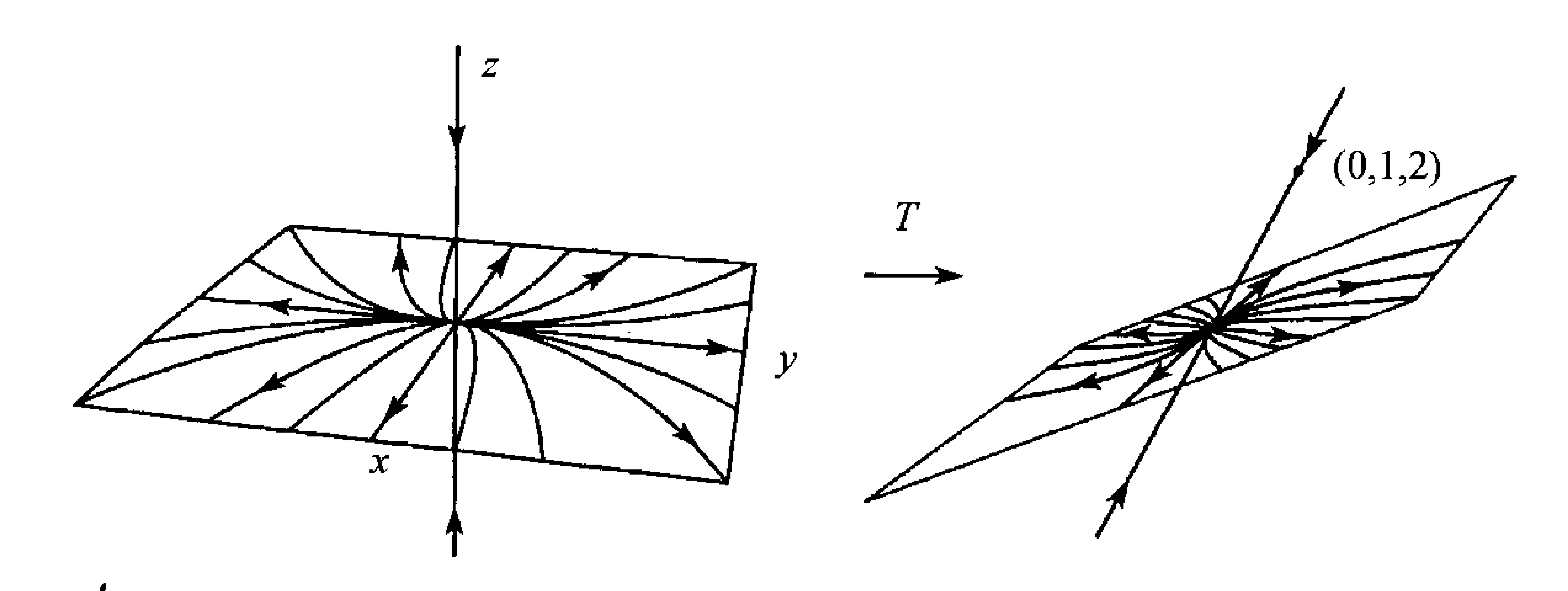


图 6.1 一个三维鞍点的稳定和不稳定子空间，左图中的系统为标准形

例 如果 3×3 矩阵 $\boldsymbol{A}$ 有 3 个不同的负特征值，则我们可作坐标变换使得系统化成

$$\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}\boldsymbol{Y},$$

其中 $\lambda_3 < \lambda_2 < \lambda_1 < 0$. 系统的所有解都趋向原点，这样我们就得到了一个高维汇点(见图 6.2). 当初值条件 (x_0, y_0, z_0) 中的 3 个坐标都非零时，对应的解将沿切于 x 轴的方向趋于原点 (见本章的习题 2). ■

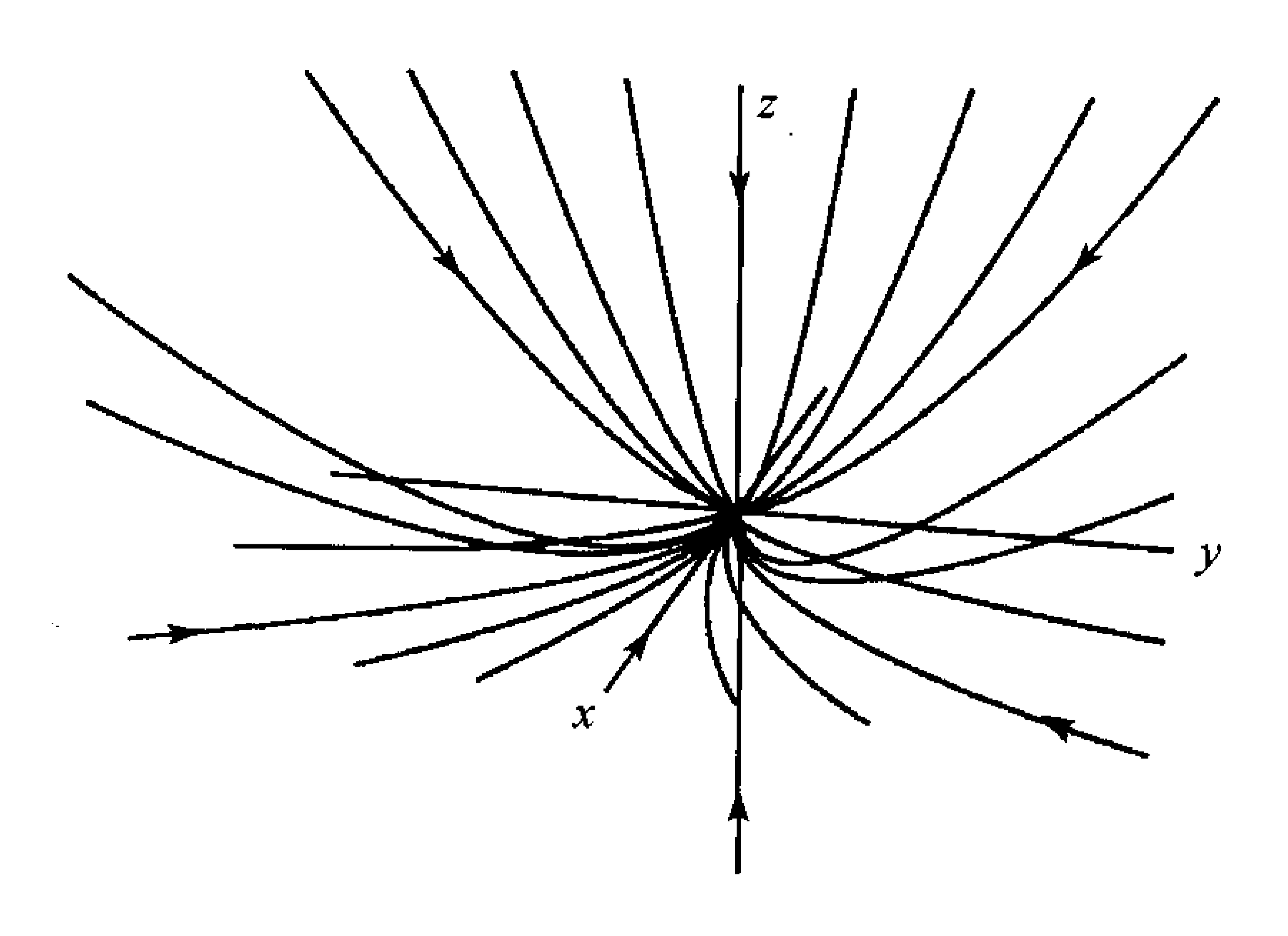


图 6.2 一个三维汇点

现在，假设 $n\times n$ 矩阵 $\boldsymbol{A}$ 具有 n 个不同特征值，其中 k_1 个是实的，$2k_2$ 个是非实的，因而 $n = k_1 + 2k_2$. 于是，由第 5 章可知，存在坐标变换使得系统化成

$$\begin{aligned} x_j' &= \lambda_j x_j \\ u_l' &= \alpha_l u_l + \beta_l v_l \\ v_l' &= -\beta_l u_l + \alpha_l v_l, \end{aligned}$$

这里 $j=1,\cdots,k_1, l=1,\cdots,k_2$. 由第 3 章可知, 我们有如下的解:

$$
\begin{aligned}
x_j(t) &= c_j\mathrm{e}^{\lambda_j t}\\
u_l(t) &= p_l\mathrm{e}^{\alpha_l t}\cos\beta_l t + q_l\mathrm{e}^{\alpha_l t}\sin\beta_l t\\
v_l(t) &= -p_l\mathrm{e}^{\alpha_l t}\sin\beta_l t + q_l\mathrm{e}^{\alpha_l t}\cos\beta_l t.
\end{aligned}
$$

像前面一样, 我们可以直接验证这就是通解. 于是, 我们就证明了:

定理 考虑系统 $\boldsymbol{X}'=\boldsymbol{AX}$, 其中 $\boldsymbol{A}$ 具有不同的特征值, $\lambda_1,\cdots,\lambda_{k_1}\in\mathbb{R}$ 而 $\alpha_1+\mathrm{i}\beta_1,\cdots,\alpha_{k_2}+\mathrm{i}\beta_{k_2}\in\mathbb{C}$. 假设 $\boldsymbol{T}$ 将 $\boldsymbol{A}$ 化为标准形

$$
\boldsymbol{T}^{-1}\boldsymbol{AT}=\begin{pmatrix}
\lambda_1 & & & & & \\
 & \ddots & & & & \\
 & & \lambda_{k_1} & & & \\
 & & & \boldsymbol{B}_1 & & \\
 & & & & \ddots & \\
 & & & & & \boldsymbol{B}_{k_2}
\end{pmatrix},
$$

其中

$$
\boldsymbol{B}_j=\begin{pmatrix}\alpha_j & \beta_j\\ -\beta_j & \alpha_j\end{pmatrix}.
$$

则 $\boldsymbol{X}'=\boldsymbol{AX}$ 的通解为 $\boldsymbol{TY}(t)$, 而

$$
\boldsymbol{Y}(t)=\begin{pmatrix}
c_1\mathrm{e}^{\lambda_1 t}\\
\vdots\\
c_{k_1}\mathrm{e}^{\lambda_{k_1} t}\\
a_1\mathrm{e}^{\alpha_1 t}\cos\beta_1 t+b_1\mathrm{e}^{\alpha_1 t}\sin\beta_1 t\\
-a_1\mathrm{e}^{\alpha_1 t}\sin\beta_1 t+b_1\mathrm{e}^{\alpha_1 t}\cos\beta_1 t\\
\vdots\\
a_{k_2}\mathrm{e}^{\alpha_{k_2} t}\cos\beta_{k_2} t+b_{k_2}\mathrm{e}^{\alpha_{k_2} t}\sin\beta_{k_2} t\\
-a_{k_2}\mathrm{e}^{\alpha_{k_2} t}\sin\beta_{k_2} t+b_{k_2}\mathrm{e}^{\alpha_{k_2} t}\cos\beta_{k_2} t
\end{pmatrix}.
$$

■

与通常一样, 矩阵 $\boldsymbol{T}$ 的列向量由对应于每个特征值的特征向量 (或特征向量的实部和虚部) 构成. 依然与前面一样, 具有负实部 (正实部) 的特征值对应的特征向量张成的子空间为稳定 (不稳定) 子空间.

例 考虑系统

$$
\boldsymbol{X}'=\begin{pmatrix}0 & 1 & 0\\ -1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}\boldsymbol{X},
$$

这里的矩阵已经是标准形, 其特征值为 $\pm\mathrm{i}, -1$. 满足初值条件 (x_0, y_0, z_0) 的解为

$$\boldsymbol{Y}(t) = x_0\begin{pmatrix}\cos t\\ -\sin t\\ 0\end{pmatrix} + y_0\begin{pmatrix}\sin t\\ \cos t\\ 0\end{pmatrix} + z_0\mathrm{e}^{-t}\begin{pmatrix}0\\ 0\\ 1\end{pmatrix},$$

这就是系统的通解. 系统的相图如图 6.3 所示. 稳定线为沿 z 轴的直线, 而 xy 平面上的所有解都绕中心为原点的圆周旋转. 事实上, 其他不在稳定线上的每一个解都位于由 "$x^2 + y^2 =$ 常数" 所确定的 $\mathbb{R}^3$ 中的圆柱面上, 当初值条件 (x_0, y_0, z_0) 中的 $z_0 \neq 0$ 时, 这些解都盘旋地趋向 xy 平面上以 $\sqrt{x_0^2 + y_0^2}$ 为半径的圆形解. ■

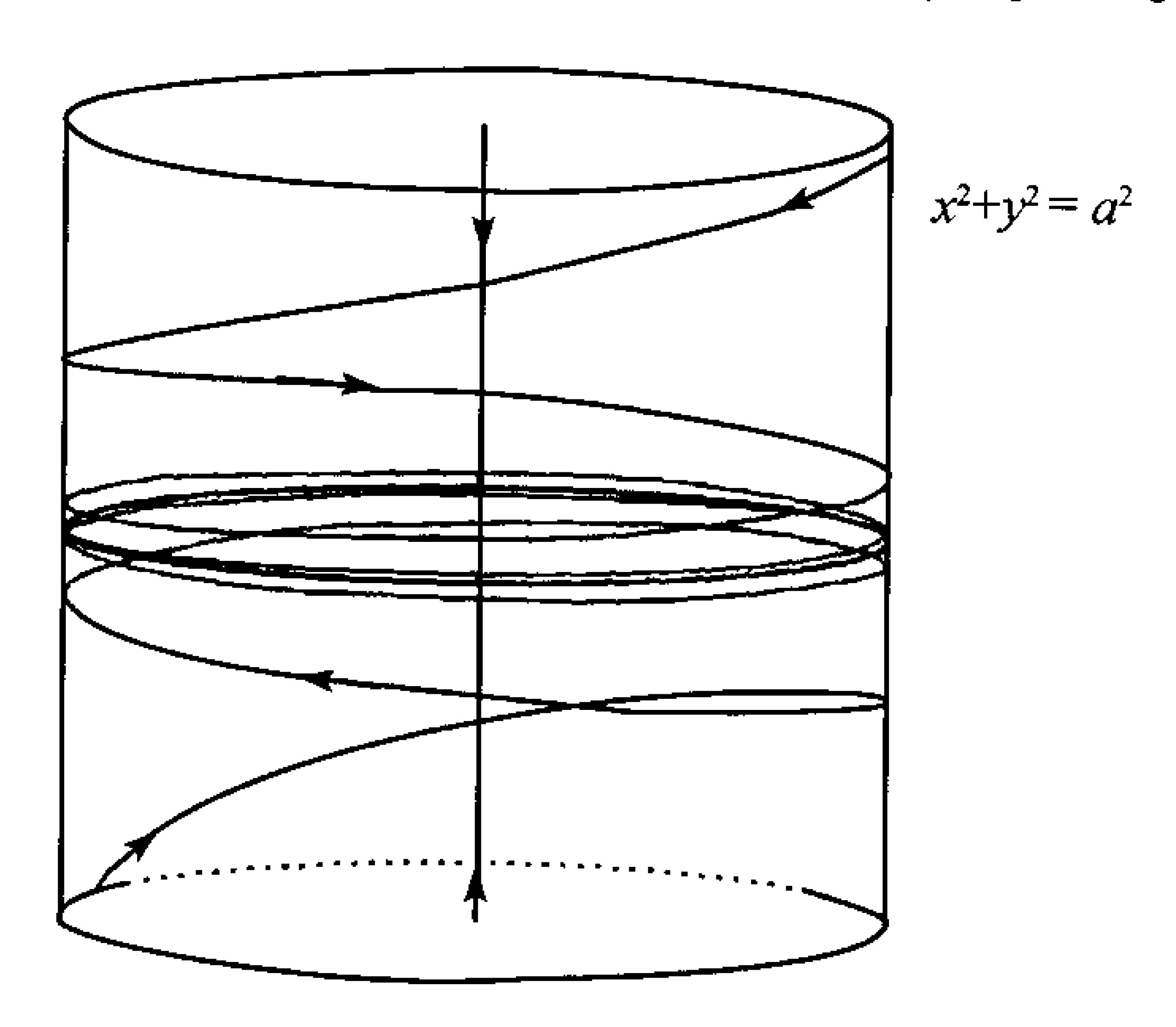


图 6.3 一个螺线中心的相图

例 现在考虑 $\boldsymbol{X}' = \boldsymbol{AX}$, 其中

$$\boldsymbol{A} = \begin{pmatrix}-0.1 & 0 & 1\\ -1 & 1 & -1.1\\ -1 & 0 & -0.1\end{pmatrix}.$$

显然它的特征方程为

$$-\lambda^3 + 0.8\lambda^2 - 0.81\lambda + 1.01 = 0,$$

我们可将它分解成

$$(1-\lambda)(\lambda^2 + 0.2\lambda + 1.01) = 0.$$

于是特征值就是上述这个方程的根, 它们是 $1, -0.1 \pm \mathrm{i}$. 求解方程 $(\boldsymbol{A} - (-0.1 + \mathrm{i})\boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 可得属于 $-0.1 + \mathrm{i}$ 的特征向量 $(-\mathrm{i}, 1, 1)$. 记 $\boldsymbol{V}_1 = \mathrm{Re}\,(-\mathrm{i}, 1, 1) = (0, 1, 1), \boldsymbol{V}_2 = \mathrm{Im}\,(-\mathrm{i}, 1, 1) = (-1, 0, 0)$. 再求解方程 $(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{X} = \boldsymbol{0}$ 可得属于 $\lambda = 1$ 的一个特征向量 $\boldsymbol{V}_3 = (0, 1, 0)$. 从而以 $\boldsymbol{V}_i$ 为列向量的矩阵

$$T=\begin{pmatrix}0 & -1 & 0\\ 1 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}$$

就将 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 化为标准形

$$\boldsymbol{Y}'=\begin{pmatrix}-0.1 & 1 & 0\\ -1 & -0.1 & 0\\ 0 & 0 & 1\end{pmatrix}\boldsymbol{Y}.$$

这个系统的不稳定线为 z 轴, 而 xy 平面为稳定平面. 注意, 稳定平面上的所有解都盘旋进入原点. 我们称这个系统为一个**螺线鞍点**(见图 6.4). 稳定平面外的典型解都将盘旋趋向于 z 轴, 而其 z 坐标的绝对值将增加或减少 (见图 6.5). ■

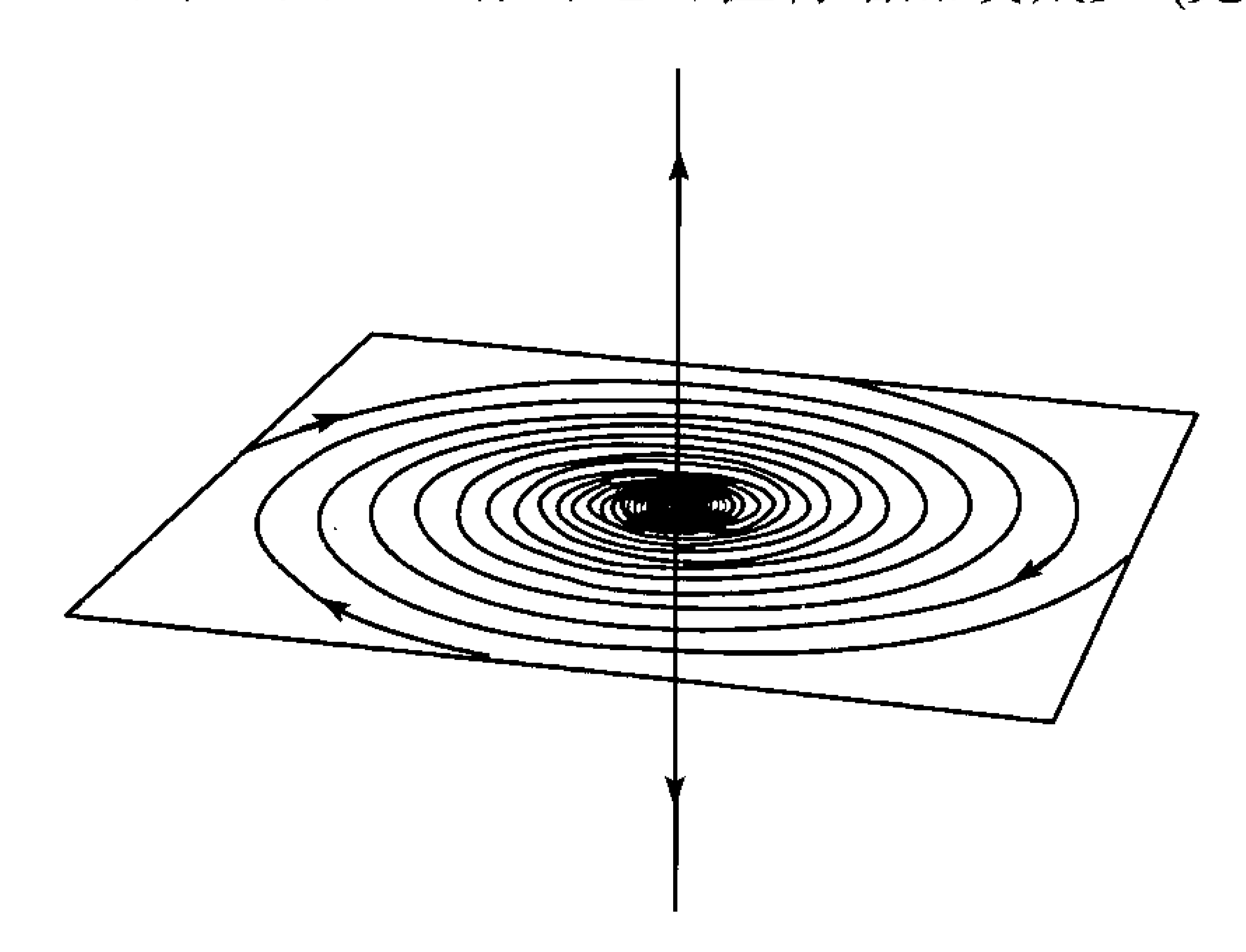


图 6.4 一个标准形式的螺线鞍点

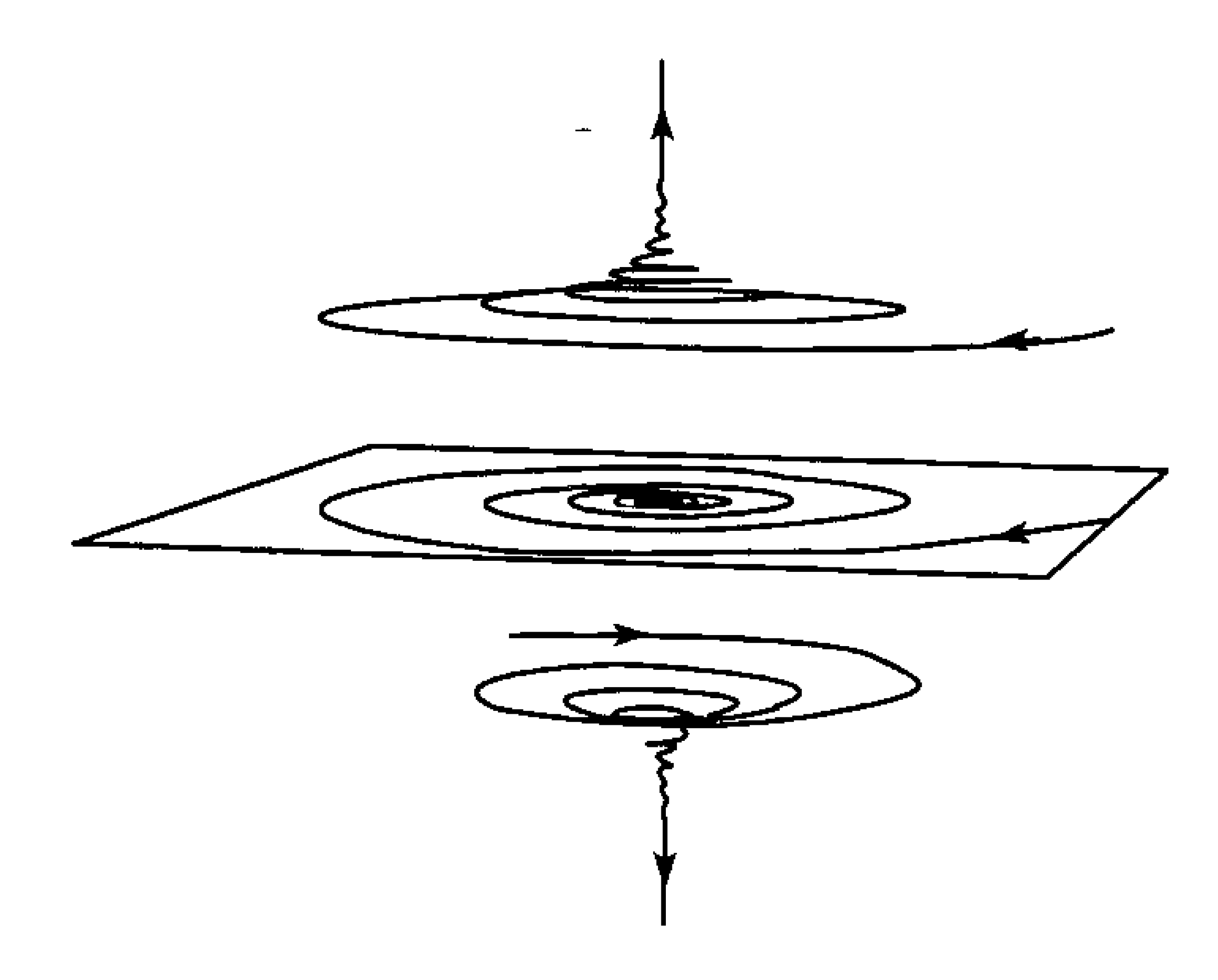


图 6.5 螺线鞍点的典型解盘旋趋向于不稳定线

6.2 调和振子

考虑一对无阻尼的调和振子, 它们的方程是:

$$x_1'' = -\omega_1^2 x_1$$

$$x_2'' = -\omega_2^2 x_2.$$

通过检验脑海中的 $\sin\omega t, \cos\omega t$ 这样的函数, 我们就几乎可以解出这两个方程. 然而, 让我们再进一步, 首先对上一节中的定理在非实特征值情形做些解释, 但更重要的是引入一些有趣的几何.

先引入两个新变量 $y_j = x_j', j = 1, 2$, 这样方程组就可以改写成下面的系统

$$x_j' = y_j$$

$$y_j' = -\omega_j^2 x_j.$$

该系统的矩阵形式为 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 其中 $\boldsymbol{X} = (x_1, y_1, x_2, y_2)$,

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & & \\ -\omega_1^2 & 0 & & \\ & & 0 & 1 \\ & & -\omega_2^2 & 0 \end{pmatrix}.$$

这个系统的特征值为 $\pm\mathrm{i}\omega_1, \pm\mathrm{i}\omega_2$. 属于 $\mathrm{i}\omega_1, \mathrm{i}\omega_2$ 的特征向量分别为 $\boldsymbol{V}_1 = (1, \mathrm{i}\omega_1, 0, 0)$, $\boldsymbol{V}_2 = (0, 0, 1, \mathrm{i}\omega_2)$. 记 $\boldsymbol{W}_1, \boldsymbol{W}_2$ 分别为 $\boldsymbol{V}_1$ 的实部和虚部, $\boldsymbol{W}_3, \boldsymbol{W}_4$ 分别为 $\boldsymbol{V}_2$ 的实部和虚部. 像通常一样, 令 $\boldsymbol{T}\boldsymbol{E}_j = \boldsymbol{W}_j$, 则线性映射 $\boldsymbol{T}$ 就将系统化为标准形, 对应的矩阵为

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{pmatrix} 0 & \omega_1 & & \\ -\omega_1 & 0 & & \\ & & 0 & \omega_2 \\ & & -\omega_2 & 0 \end{pmatrix}.$$

这样就可以得到通解

$$\boldsymbol{Y}(t) = \begin{pmatrix} x_1(t) \\ y_1(t) \\ x_2(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} a_1\cos\omega_1 t + b_1\sin\omega_1 t \\ -a_1\sin\omega_1 t + b_1\cos\omega_1 t \\ a_2\cos\omega_2 t + b_2\sin\omega_2 t \\ -a_2\sin\omega_2 t + b_2\cos\omega_2 t \end{pmatrix},$$

它的形式正像我们所想像的那样.

我们也许会认为事情就到此为止了, 因为解的表达式都已经得到了, 但是, 我们将往前更进一步.

显然, 每对解 $(x_j(t), y_j(t)), j = 1, 2$ 都是单个方程的一个以 $2\pi/\omega_j$ 为周期的周期解, 但这并不意味着完整的四维解也是一个周期解. 事实上, 完整的解是一个周期为 τ 的周期函数当且仅当存在整数 m, n 使得

$$\omega_1\tau = m \cdot 2\pi, \qquad \omega_2\tau = n \cdot 2\pi.$$

从而, 为了得到周期性, 我们必须有

$$\tau = \frac{2\pi m}{\omega_1} = \frac{2\pi n}{\omega_2},$$

这等价于

$$\frac{\omega_2}{\omega_1} = \frac{n}{m}.$$

即, 调和振子的两个频率比必须是一个有理数. 在图 6.6 中, 我们画出了当频率之比为 5/2 时系统的一个特解 $(x_1(t), x_2(t))$.

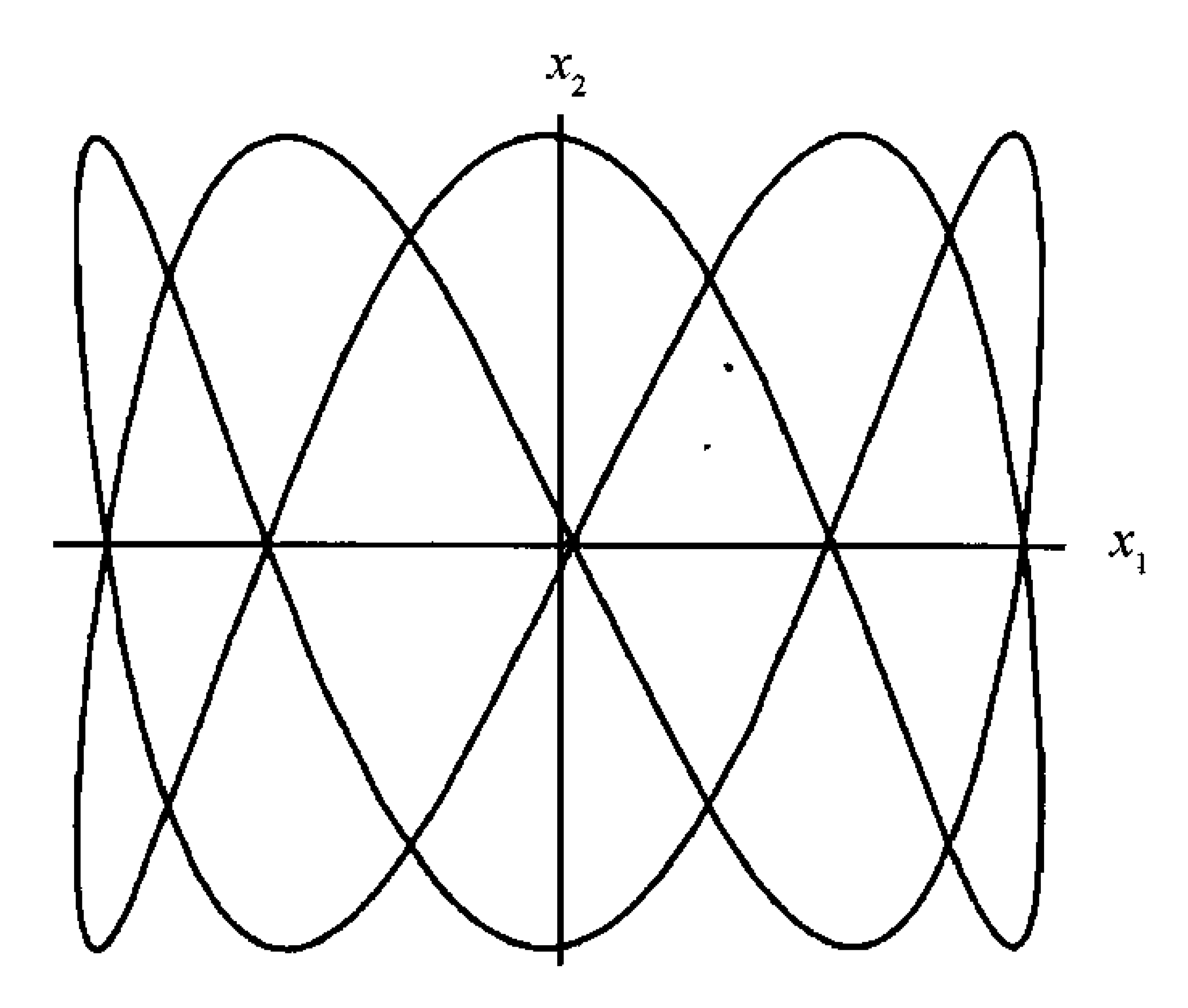


图 6.6　频率之比为 5/2 时一个解在 x_1x_2 平面上的投影. 可见在回到初始位置之前, $x_2(t)$ 要作 5 次振动, 而 $x_1(t)$ 只作 2 次

当频率之比为无理数时, 情况非常不一样. 为了理解这点, 我们作另外一个 (大家非常熟悉的) 坐标变换. 用标准形来写, 现在的系统是

$$x_j' = w_j y_j$$

$$y_j' = -\omega_j x_j.$$

我们引入极坐标 (r_j, θ_j) 来代替变量 x_j, y_j. 对式子

$$r_j^2 = x_j^2 + y_j^2$$

两边求导可得

$$2r_jr_j' = 2x_jx_j' + 2y_jy_j' = 2x_jy_j\omega_j - 2x_jy_j\omega_j = 0.$$

于是,对每一个 j 都有 $r_j' = 0$. 同样地, 对方程

$$\tan\theta_j = \frac{y_j}{x_j}$$

求导可得

$$(\sec^2\theta_j)\theta_j' = \frac{y_j'x_j - y_jx_j'}{x_j^2} = \frac{-\omega_j r_j^2}{r_j^2\cos^2\theta_j},$$

由此可得

$$\theta_j' = -\omega_j.$$

从而在极坐标下, 这些方程就变得非常简单:

$$r_j' = 0$$

$$\theta_j' = -\omega_j.$$

第一个方程告诉我们, r_1, r_2 沿任何解都始终为常数. 从而, 无论我们如何选取 r_1, r_2 的初始值, 关于 θ_j 的方程都是一样的. 这样我们就可以将我们的注意力集中到 $r_1 = r_2 = 1$ 的情形. 在 $\mathbb{R}^4$ 中得到的点集是一个环面, 也就是油炸圈饼的表面. 虽然在四维空间中这有点难以想象, 但不管怎样, 这个集合上有 2 个独立的变量, 即 θ_1, θ_2, 而且它们都是以 2π 为周期的. 在这一点上, 这类似于 $\mathbb{R}^3$ 中熟知的环面可以被 2 个独立的循环走向所参数化.

现在, 方程限制在环面上就变成了:

$$\theta_1' = -\omega_1$$

$$\theta_2' = -\omega_2.$$

为了方便, 我们可将 θ_1, θ_2 看成是边长为 2π 的正方形中的变量 (只要将正方形的对边 $\theta_j = 0$ 和 $\theta_j = 2\pi$ 粘合起来就得到了环面). 在正方形中, 向量场的斜率为常数

$$\frac{\theta_2'}{\theta_1'} = \frac{\omega_2}{\omega_1}.$$

于是, 在正方形中, 解都位于斜率为 ω_2/ω_1 的直线上. 当解到达 $\theta_1 = 2\pi$ 这条边时 (不妨设在 $\theta_2 = c$ 处), 它就在 $\theta_1 = 0$ 的边上马上重新出现 (θ_2 的坐标依然为 c), 然后以斜率 ω_2/ω_1 再继续向前. 在解达到 $\theta_2 = 2\pi$ 时, 将会出现类似的等同.

这样, 我们就对解在环面上的行为有了一个简化的几何图像. 将会出现些什么呢? 答案依赖于 ω_2/ω_1 的取值. 如果这个值为有理数, 比如, n/m, 则解从 $(\theta_1(0), \theta_2(0))$ 出发将垂直经过环面 n 次, 水平经过环面 m 次之后再回到出发点. 这就是我们上面观察到的周期解. 有时, 直线解在 $\theta_1\theta_2$ 平面上的图像和图 6.6 中所画出的 x_1x_2 平面上的图像并不完全一样.

在无理情形, 将会出现一些很不同的情况 (见图 6.7). 为了理解现在出现的情况, 让我们回到第 1 章讨论过的**庞加莱映射**的概念. 考虑圆周 $\theta_1 = 0$, 也就是环面正方形表示中的左边. 任给该圆周上的一个初值, 比如, $\theta_2 = x_0$, 我们跟随从该点出发的解一直到它下次再达到 $\theta_1 = 2\pi$. 由于环面是等同正方形的对边得到的, 解现在实际上就是回到了出发的圆周 $\theta_1 = 0$. 在这个行进过程中, 解有可能会穿过边界 $\theta_2 = 2\pi$ 几次, 但它最终会回到 $\theta_1 = 0$. 因而我们可以定义 $\theta_1 = 0$ 上的庞加莱映射如下: 庞加莱映射在圆周上 x_0 处的取值就是首次返回点的相应坐标. 假设首次返回出现在点 $\theta_2(\tau)$ 处, 其中 τ 是满足 $\theta_1(\tau) = 2\pi$ 的时刻. 由于 $\theta_1(t) = \theta_1(0) + \omega_1 t$, 因而 $\tau = 2\pi/\omega_1$. 从而 $\theta_2(\tau) = x_0 + \omega_2(2\pi/\omega_1)$. 于是在这个圆周上的庞加莱映射可以写成

$$f(x_0) = x_0 + 2\pi(\omega_2/\omega_1) \bmod 2\pi,$$

其中 $x_0 = \theta_2(0)$ 是初值在这个圆周上的 θ_2 坐标 (见图 6.8). 于是, 该圆周上的庞加莱映射就是将圆周上的点旋转角度 $2\pi(\omega_2/\omega_1)$ 这个函数. 由于 ω_2/ω_1 是无理的, 这个函数称为圆周上的**无理旋转**.

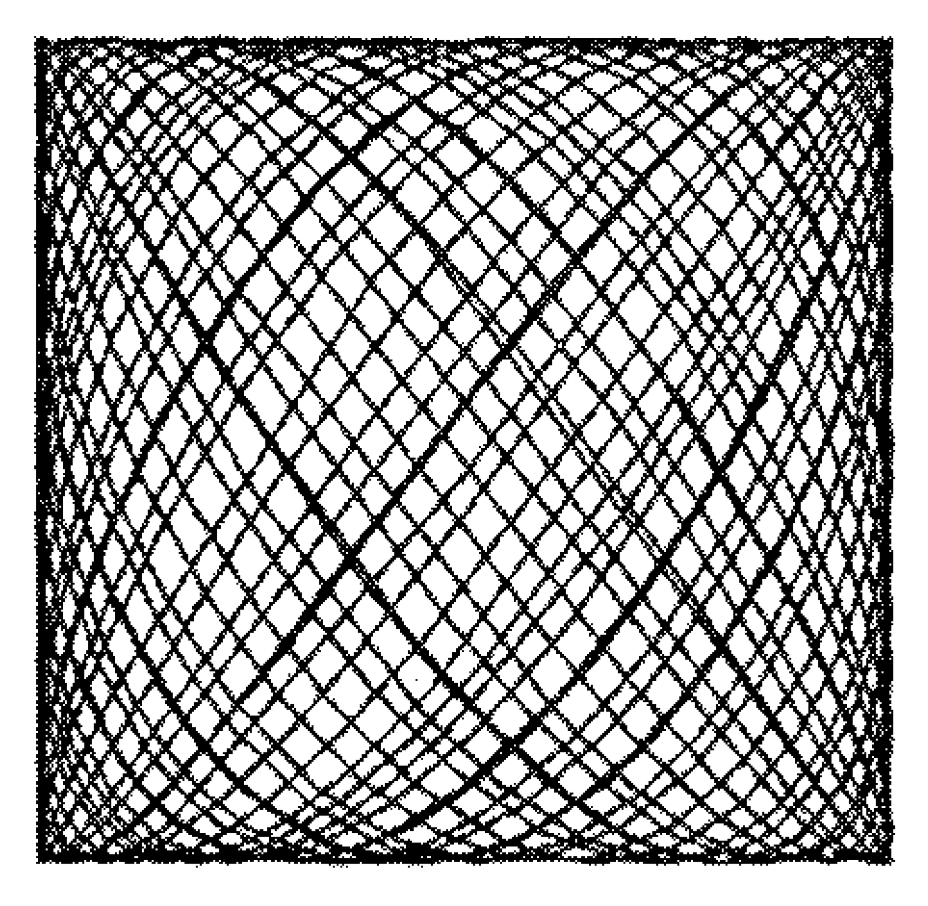
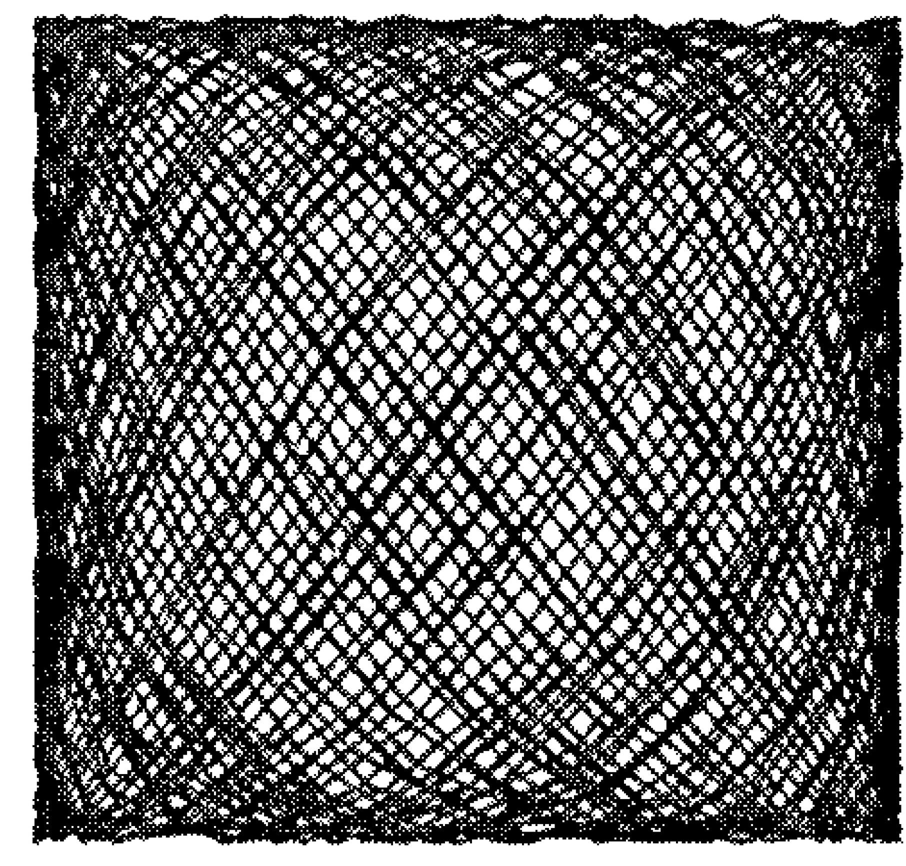



图 6.7 频率之比为 $\sqrt{2}$ 时一个解在 x_1x_2 平面上的投影, 左边的曲线计算到时间 50π, 右边的到 100π

定义 点集 $x_0, x_1 = f(x_0), x_2 = f(f(x_0)), \cdots, x_n = f(x_{n-1})$ 称为 x_0 在 f 迭代下的**轨道**. ■

x_0 的轨道记录了当时间增加时, 解是如何穿过 $\theta_1 = 2\pi$ 的.

命题 假设 ω_2/ω_1 是无理数, 则圆周 $\theta_1 = 0$ 上任何初值 x_0 的轨道在该圆周上都是稠密的.

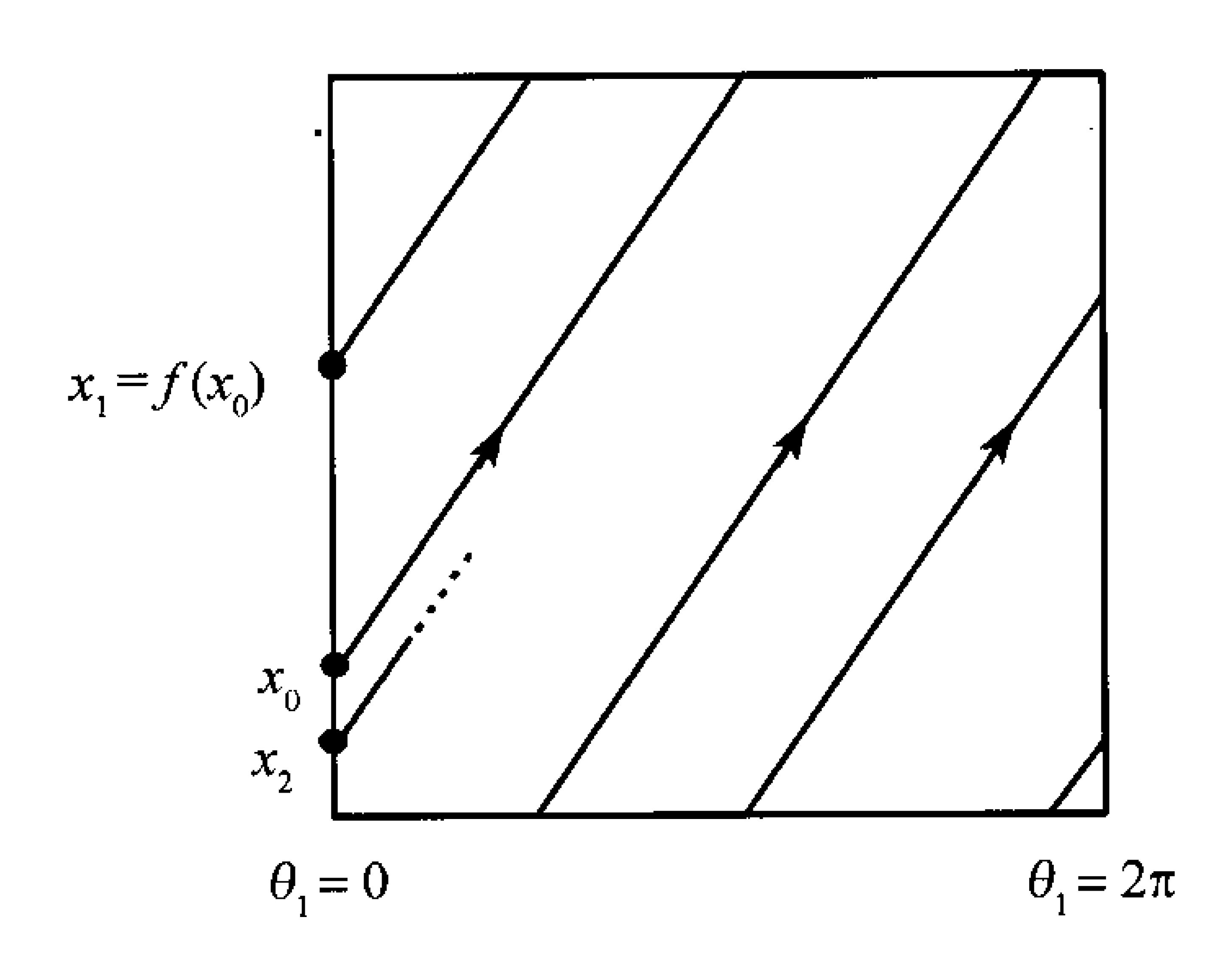


图 6.8 在 $\theta_1\theta_2$ 环面上, 关于圆周 $\theta_1=0$ 的庞加莱映射[①]

证明 由 5.6 节中的定义, 圆周的一个子集是**稠密的**, 如果对圆周上的任一点, 这个子集中都有点去任意接近. 因而, 我们必须证明: 任给圆周上的点 z, 以及任意的 $\epsilon>0$, 总存在 x_0 轨道上的一点 x_n, 使得 $|z-x_n|<\epsilon$, 其中 z 和 x_n 之间的距离是在 mod 2π 意义下的. 首先观察到, 总存在 $n,m,m>n$, 满足 $|x_n-x_m|<\epsilon$. 事实上, 由于 ω_2/ω_1 是无理的, x_0 的轨道不是有限点集. 又由于圆周的周长有限, 因而在这些点中必定至少存在两个点使得它们的距离小于 ϵ. 这就是要找的点 x_n,x_m (实际上, 有无限多个这样的点). 现在将这对点反向旋转 n 次. 于是, x_n,x_m 将分别旋转到 x_0,x_{m-n}. 这样, 经过旋转之后, 就有 $|x_0-x_{m-n}|<\epsilon$. 从而 x_{m-n} 是通过将圆周旋转角度 $(m-n)2\pi(\omega_2/\omega_1)$ 得到的 (在 mod 2π 后, 这是一个角度小于 ϵ 的旋转). 因而, 如果再作这个旋转, 依然还有

$$|x_{2(m-n)}-x_{m-n}|<\epsilon,$$

归纳地就可得到, 对所有的 k,

$$|x_{k(m-n)}-x_{(k-1)(m-n)}|<\epsilon.$$

这样, 通过一次次旋转角度 $(m-n)2\pi(\omega_2/\omega_1)$, 我们就得到了一列点, 其中相邻的两点之间的距离都在 ϵ 以内. 因而这其中必有一个点在 z 的 ϵ 范围以内. ■

因为 x_0 的轨道在圆周 $\theta_1=0$ 上是稠密的, 从而在正方形中将这些点连接起来的直线解也是稠密的, 于是原来的解在其所在的环面上是稠密的. 这就解释了为什么当 $\omega_2/\omega_1=\sqrt{2}$ 时, 解如图 6.7 所示投影到 x_1x_2 平面时会稠密地布满.

回到振子的实际运动, 我们就得到当 ω_2/ω_1 为无理时, 两个质点的运动并不是周期的. 然而, 由于解在环面上的稠密性, 当时间增加时, 它们又的确会一次次地回到非常靠近初始位置的地方. 这种形式的运动称为**拟周期运动**. 在习题 7 中, 我们还要考察一类相互关联的方程组, 即一对耦合振子.

① 图中的箭头应该对应于负的时间方向.—— 译者注

6.3 重特征值

在上一章中，我们已经看到，具有实重特征值系统的求解归结为求解其矩阵含有如下形式的块的系统：

$$\begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ & & & & \lambda \end{pmatrix}.$$

例 令

$$\boldsymbol{X}' = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \boldsymbol{X}.$$

这个系统唯一的特征值就是 λ，而且唯一的特征向量为 $(1,0,0)$. 我们像在第 3 章所做的那样来求解这个系统. 首先，由于 $x_3' = \lambda x_3$，故必有

$$x_3(t) = c_3 e^{\lambda t},$$

进而有，

$$x_2' = \lambda x_2 + c_3 e^{\lambda t}.$$

与第 3 章一样，我们猜测它有如下形式的解：

$$x_2(t) = c_2 e^{\lambda t} + \alpha t e^{\lambda t}.$$

将上式代入 x_2' 的微分方程，可以定出 $\alpha = c_3$，从而找到

$$x_2(t) = c_2 e^{\lambda t} + c_3 t e^{\lambda t}.$$

最后，方程

$$x_1' = \lambda x_1 + c_2 e^{\lambda t} + c_3 t e^{\lambda t}$$

让人联想起下面形式的解：

$$x_1(t) = c_1 e^{\lambda t} + \alpha t e^{\lambda t} + \beta t^2 e^{\lambda t}.$$

像刚才一样的求解可得

$$x_1(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t} + c_3 \frac{t^2}{2} e^{\lambda t}.$$

总之, 我们找到了

$$X(t)=c_1\mathrm{e}^{\lambda t}\begin{pmatrix}1\\0\\0\end{pmatrix}+c_2\mathrm{e}^{\lambda t}\begin{pmatrix}t\\1\\0\end{pmatrix}+c_3\mathrm{e}^{\lambda t}\begin{pmatrix}t^2/2\\t\\1\end{pmatrix},$$

它就是系统的通解. 虽然解的表达式中出现了多项式项, 但当 $\lambda<0$ 时, 指数项将起主导作用, 而且所有的解都趋于零. 当 $\lambda<0$ 时, 一些有代表性的解如图 6.9 所示. 注意, 该系统只有唯一的直线解, 该解位于 x 轴上, 而且 xy 平面是不变的, 其上解的形态与平面重特征值情形完全相同. ■

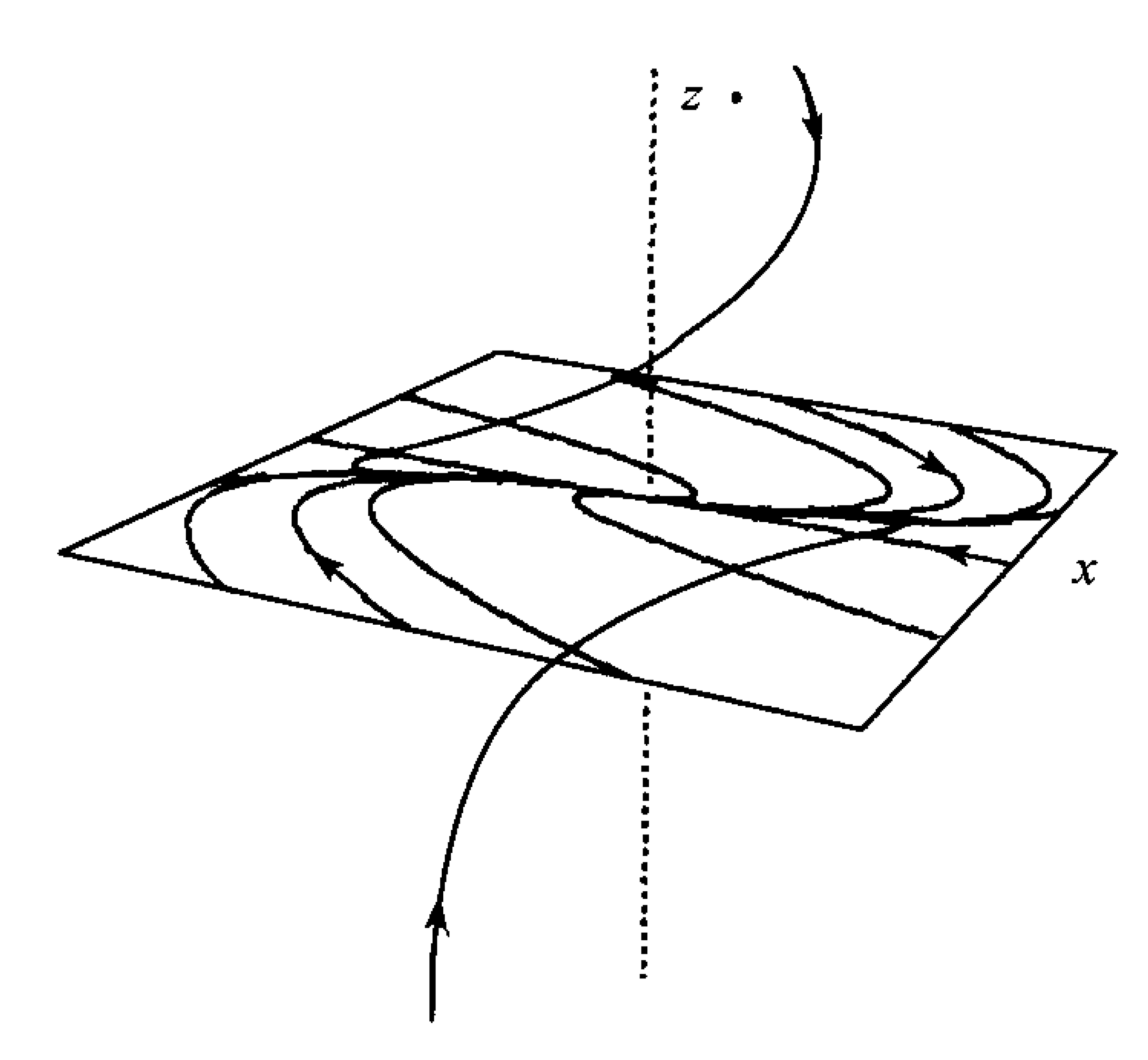


图 6.9 重实特征值情形的相图

例 考虑下面的四维系统:

$$x_1'=x_1+x_2-x_3$$

$$x_2'=x_2+x_4$$

$$x_3'=x_3+x_4$$

$$x_4'=x_4$$

我们可将该系统写成如下的矩阵形式:

$$X'=AX=\begin{pmatrix}1&1&-1&0\\0&1&0&1\\0&0&1&1\\0&0&0&1\end{pmatrix}X.$$

因为 A 为上三角的, 故它的所有的特征值都是 1. 求解 $(A-I)X=0$ 可找到两

个线性无关的特征向量 $\boldsymbol{V}_1 = (1,0,0,0), \boldsymbol{W}_1 = (0,1,1,0)$. 这就将 $\boldsymbol{A}$ 可能的标准形减少到两种. 进一步求解 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{X} = \boldsymbol{V}_1$ 可得一个解 $\boldsymbol{V}_2 = (0,1,0,0)$, 求解 $(\boldsymbol{A}-\boldsymbol{I})\boldsymbol{X} = \boldsymbol{W}_1$ 可得一个解 $\boldsymbol{W}_2 = (0,0,0,1)$. 由此, 我们知道系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 可以变换成

$$\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \boldsymbol{Y},$$

其中矩阵 $\boldsymbol{T}$ 由下式给出:

$$\boldsymbol{T} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

于是 $\boldsymbol{Y}' = (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})\boldsymbol{Y}$ 的解为

$$\begin{aligned} y_1(t) &= c_1 \mathrm{e}^t + c_2 t\mathrm{e}^t \\ y_2(t) &= c_2 \mathrm{e}^t \\ y_3(t) &= c_3 \mathrm{e}^t + c_4 t\mathrm{e}^t \\ y_4(t) &= c_4 \mathrm{e}^t. \end{aligned}$$

再用坐标变换 $\boldsymbol{T}$ 作用, 我们就得到了原系统的通解为

$$\begin{aligned} x_1(t) &= c_1 \mathrm{e}^t + c_2 t\mathrm{e}^t \\ x_2(t) &= c_2 \mathrm{e}^t + c_3 \mathrm{e}^t + c_4 t\mathrm{e}^t \\ x_3(t) &= c_3 \mathrm{e}^t + c_4 t\mathrm{e}^t \\ x_4(t) &= c_4 \mathrm{e}^t. \end{aligned}$$

■

6.4 矩阵指数

现在, 我们转到用矩阵指数求解线性系统, 这是解线性系统的一个不同但简练的方法. 从某种意义上讲, 这是攻克这类系统更加自然的方法.

我们先回忆 1.1 节中是如何求解线性方程 $x' = ax$ 的 1×1 “系统”, 其中我们的矩阵就是简单的 (a). 那时我们并没有进行找特征值和特征向量的过程, (然而,

事实上,我们进行了,但过程实在是太简单了.)而是直接取矩阵 (a) 的指数函数就找到了通解 $x(t)=c\exp(at)$. 事实上,这个过程对一般 $n\times n$ 矩阵 $\boldsymbol{A}$ 也是行得通的. 我们唯一需要知道的就是如何求一个矩阵的指数.

办法如下. 由微积分可知,指数函数可以表示成一个无穷级数

$$\mathrm{e}^x=\sum_{k=0}^{\infty}\frac{x^k}{k!},$$

而且我们知道对每个 $x\in\mathbb{R}$ 这个级数都收敛. 由于我们可以将矩阵相加,可以取它们的 k 次方,还可以将矩阵的每个元素乘上 $1/k!$,这将意味着我们一样可以用上面的级数去取它们的指数.

定义 设 $\boldsymbol{A}$ 为一 $n\times n$ 矩阵. 我们定义 $\boldsymbol{A}$ 的**指数**是由下式给出的矩阵:

$$\exp(\boldsymbol{A})=\sum_{k=0}^{\infty}\frac{\boldsymbol{A}^k}{k!}.$$ ■

当然,我们还要担心上面矩阵和式收敛的含义. 但把这个问题放下,我们先计算几个例子.

例 令

$$\boldsymbol{A}=\begin{pmatrix}\lambda & 0\\ 0 & \mu\end{pmatrix}.$$

则我们有

$$\boldsymbol{A}^k=\begin{pmatrix}\lambda^k & 0\\ 0 & \mu^k\end{pmatrix},$$

于是

$$\exp(\boldsymbol{A})=\begin{pmatrix}\sum_{k=0}^{\infty}\lambda^k/k! & 0\\ 0 & \sum_{k=0}^{\infty}\mu^k/k!\end{pmatrix}=\begin{pmatrix}\mathrm{e}^\lambda & 0\\ 0 & \mathrm{e}^\mu\end{pmatrix},$$

这个结果大家可能已经事先猜到. ■

例 下面来看一个稍微复杂些的例子. 令

$$\boldsymbol{A}=\begin{pmatrix}0 & \beta\\ -\beta & 0\end{pmatrix}.$$

由计算可得

$$\boldsymbol{A}^0=\boldsymbol{I},\quad \boldsymbol{A}^2=-\beta^2\boldsymbol{I},\quad \boldsymbol{A}^3=-\beta^3\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix},$$

$$A^4=\beta^4 I,\quad A^5=\beta^5\begin{pmatrix}0&1\\-1&0\end{pmatrix},\cdots,$$

于是就有

$$\exp(A)=\begin{pmatrix}\sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k}}{(2k)!} & \sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k+1}}{(2k+1)!}\\ -\sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k+1}}{(2k+1)!} & \sum_{k=0}^{\infty}(-1)^k\frac{\beta^{2k}}{(2k)!}\end{pmatrix}=\begin{pmatrix}\cos\beta & \sin\beta\\ -\sin\beta & \cos\beta\end{pmatrix}.$$

■

例 现在令

$$A=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix},$$

其中 $\lambda\neq 0$. 为了照顾随后的应用, 我们来计算 $\exp(tA)$, 而不是 $\exp A$. 我们有

$$(tA)^k=\begin{pmatrix}(t\lambda)^k & kt^k\lambda^{k-1}\\ 0 & (t\lambda)^k\end{pmatrix}.$$

于是我们得到

$$\exp(tA)=\begin{pmatrix}\sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!} & t\sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!}\\ 0 & \sum_{k=0}^{\infty}\frac{(t\lambda)^k}{k!}\end{pmatrix}=\begin{pmatrix}e^{t\lambda} & te^{t\lambda}\\ 0 & e^{t\lambda}\end{pmatrix}.$$

■

可以看到, 在上面 3 个例子中, 矩阵 $\exp(A)$ 的元素都是一些无穷级数. 因此, 我们称矩阵无穷级数 $\exp(A)$ 绝对收敛, 如果每一个单独的项绝对收敛. 在上面几种情形, 收敛性是显然的. 遗憾的是, 对于一般的矩阵 A, 收敛性并不清楚. 为了证明这里的收敛性, 我们需要做一些努力.

记 $a_{ij}(k)$ 为 A^k 的 ij 元素. 令 $a=\max|a_{ij}|$. 我们有

$$\begin{aligned}|a_{ij}(2)| &= \left|\sum_{k=1}^{n}a_{ik}a_{kj}\right|\leqslant na^2\\ |a_{ij}(3)| &= \left|\sum_{k,l=1}^{n}a_{ik}a_{kl}a_{lj}\right|\leqslant n^2a^3\\ &\vdots\\ |a_{ij}(k)| &\leqslant n^{k-1}a^k.\end{aligned}$$

于是, 我们得到了 $n \times n$ 矩阵 $\exp(\boldsymbol{A})$ 的 ij 元素的一个上界:

$$\left|\sum_{k=0}^{\infty} \frac{a_{ij}(k)}{k!}\right| \leqslant \sum_{k=0}^{\infty} \frac{|a_{ij}(k)|}{k!} \leqslant \sum_{k=0}^{\infty} \frac{n^{k-1}a^k}{k!} \leqslant \sum_{k=0}^{\infty} \frac{(na)^k}{k!} \leqslant \exp(na),$$

因此, 根据比较判别法, 这个级数绝对收敛. 这样, 对任何 $\boldsymbol{A} \in L(\mathbb{R}^n)$, $\exp(\boldsymbol{A})$ 有意义.

下面的结果表明, 矩阵指数也具有通常指数函数的许多熟悉的性质.

命题 设 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{T}$ 都是 $n \times n$ 矩阵, 则

(1) 如果 $\boldsymbol{B} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$, 则 $\exp(\boldsymbol{B}) = \boldsymbol{T}^{-1}\exp(\boldsymbol{A})\boldsymbol{T}$.

(2) 如果 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$, 则 $\exp(\boldsymbol{A}+\boldsymbol{B}) = \exp(\boldsymbol{A})\exp(\boldsymbol{B})$.

(3) $\exp(-\boldsymbol{A}) = (\exp(\boldsymbol{A}))^{-1}$.

证明 根据恒等式 $\boldsymbol{T}^{-1}(\boldsymbol{A}+\boldsymbol{B})\boldsymbol{T} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} + \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}, (\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})^k = \boldsymbol{T}^{-1}\boldsymbol{A}^k\boldsymbol{T}$, 我们有

$$\boldsymbol{T}^{-1}\left(\sum_{k=0}^{n} \frac{\boldsymbol{A}^k}{k!}\right)\boldsymbol{T} = \sum_{k=0}^{n} \frac{(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T})^k}{k!},$$

两边取极限就得到了 (1).

为了证明 (2), 因为 $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$, 由二项式定理我们有

$$(\boldsymbol{A}+\boldsymbol{B})^n = n! \sum_{j+k=n} \frac{\boldsymbol{A}^j}{j!}\frac{\boldsymbol{B}^k}{k!}.$$

这样我们就必须证明

$$\sum_{n=0}^{\infty}\left(\sum_{j+k=n} \frac{\boldsymbol{A}^j}{j!}\frac{\boldsymbol{B}^k}{k!}\right) = \left(\sum_{j=0}^{\infty} \frac{\boldsymbol{A}^j}{j!}\right)\left(\sum_{k=0}^{\infty} \frac{\boldsymbol{B}^k}{k!}\right).$$

这并不像它看起来那么显然, 因为我们处理的是矩阵的级数而不是实数的级数. 我们将在下面的引理中证明上式, 然后就证明了 (2). 在 (2) 中令 $\boldsymbol{B} = -\boldsymbol{A}$ 就可以推出 (3). ■

引理 对于任意的 $n \times n$ 矩阵 $\boldsymbol{A}, \boldsymbol{B}$, 我们有

$$\sum_{n=0}^{\infty}\left(\sum_{j+k=n} \frac{\boldsymbol{A}^j}{j!}\frac{\boldsymbol{B}^k}{k!}\right) = \left(\sum_{j=0}^{\infty} \frac{\boldsymbol{A}^j}{j!}\right)\left(\sum_{k=0}^{\infty} \frac{\boldsymbol{B}^k}{k!}\right).$$

证明 我们已经知道这几个矩阵无穷级数都收敛, 因而只须验证它们的极限相同. 为此, 考虑部分和

$$\gamma_{2m} = \sum_{n=0}^{2m}\left(\sum_{j+k=n} \frac{\boldsymbol{A}^j}{j!}\frac{\boldsymbol{B}^k}{k!}\right)$$

和

$$\alpha_m = \left(\sum_{j=0}^{m} \frac{A^j}{j!}\right), \qquad \beta_m = \left(\sum_{k=0}^{m} \frac{B^k}{k!}\right).$$

我们需要说明当 $m \to \infty$ 时, 矩阵 $\gamma_{2m} - \alpha_m\beta_m$ 趋于零矩阵. 在下面证明中, 对于任一矩阵 $M = [m_{ij}]$, 记 $|M| = \max|m_{ij}|$. 我们将说明当 $m \to \infty$ 时, $|\gamma_{2m} - \alpha_m\beta_m|$ 趋于零.

计算表明

$$\gamma_{2m} - \alpha_m\beta_m = \sum{}' \frac{A^j}{j!}\frac{B^k}{k!} + \sum{}'' \frac{A^j}{j!}\frac{B^k}{k!},$$

其中 $\sum{}'$ 表示在指标满足

$$j + k \leqslant 2m, \quad 0 \leqslant j \leqslant m, \quad m+1 \leqslant k \leqslant 2m$$

的项上求和, 而 $\sum{}''$ 则表示在指标满足

$$j + k \leqslant 2m, \quad m+1 \leqslant j \leqslant 2m, \quad 0 \leqslant k \leqslant m$$

的项上求和. 于是

$$||\gamma_{2m} - \alpha_m\beta_m|| \leqslant \sum{}' \left\|\frac{A^j}{j!}\right\| \cdot \left\|\frac{B^k}{k!}\right\| + \sum{}'' \left\|\frac{A^j}{j!}\right\| \cdot \left\|\frac{B^k}{k!}\right\|.$$

现在又有

$$\sum{}' \left\|\frac{A^j}{j!}\right\| \cdot \left\|\frac{B^k}{k!}\right\| \leqslant \left(\sum_{j=0}^{m} \left\|\frac{A^j}{j!}\right\|\right) \left(\sum_{k=m+1}^{2m} \left\|\frac{B^k}{k!}\right\|\right).$$

当 $m \to \infty$ 时, 上式趋于 0, 这是因为在前面我们已经看到

$$\sum_{j=0}^{\infty} \left\|\frac{A^j}{j!}\right\| \leqslant \exp(n|A|) < \infty.$$

同样地, 当 $m \to \infty$ 时,

$$\sum{}'' \left\|\frac{A^j}{j!}\right\| \cdot \left\|\frac{B^k}{k!}\right\| \to 0.$$

于是, $\lim_{m\to\infty}(\gamma_{2m} - \alpha_m\beta_m) = 0$, 这就证明了引理. ■

注意, 命题中论断 (3) 意味着, 对每一个矩阵 A, $\exp(A)$ 都可逆. 这一点类似于对每个实数 a, $e^a \neq 0$.

在 $\boldsymbol{A}$ 和 $\exp(\boldsymbol{A})$ 的特征向量之间有一个很简单的关系：

命题 如果 $\boldsymbol{V} \in \mathbb{R}^n$ 是 $\boldsymbol{A}$ 的属于特征值 λ 的特征向量, 则 $\boldsymbol{V}$ 也是 $\exp(\boldsymbol{A})$ 的特征向量, 它属于 e^λ.

证明 由 $\boldsymbol{A}\boldsymbol{V} = \lambda \boldsymbol{V}$ 可得,

$$\exp(\boldsymbol{A})\boldsymbol{V} = \lim_{n\to\infty}\left(\sum_{k=0}^{n}\frac{\boldsymbol{A}^k\boldsymbol{V}}{k!}\right) = \lim_{n\to\infty}\left(\sum_{k=0}^{n}\frac{\lambda^k}{k!}\boldsymbol{V}\right) = \left(\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\right)\boldsymbol{V} = \mathrm{e}^\lambda\boldsymbol{V}. \quad \blacksquare$$

现在我们回到微分方程系统的讨论. 令 $\boldsymbol{A}$ 为一 $n\times n$ 矩阵, 考虑系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$. 回想我们以前用 $L(\mathbb{R}^n)$ 记全体 $n \times n$ 矩阵的集合. 把每个 $t \in \mathbb{R}$ 映到 $\exp(t\boldsymbol{A})$ 就得到了一个函数 $\mathbb{R} \to L(\mathbb{R}^n)$. 由于我们将 $L(\mathbb{R}^n)$ 等同于 $\mathbb{R}^{n^2}$, 因而谈论这个函数的导数是有意义的.

命题

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(t\boldsymbol{A}) = \boldsymbol{A}\exp(t\boldsymbol{A}) = \exp(t\boldsymbol{A})\boldsymbol{A}.$$

换句话说, 矩阵值函数 $t \to \exp(t\boldsymbol{A})$ 的导数是另一个矩阵值函数 $\boldsymbol{A}\exp(t\boldsymbol{A})$.

证明 我们有

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\exp(t\boldsymbol{A}) &= \lim_{h\to 0}\frac{\exp((t+h)\boldsymbol{A}) - \exp(t\boldsymbol{A})}{h}\\ &= \lim_{h\to 0}\frac{\exp(t\boldsymbol{A})\exp(h\boldsymbol{A}) - \exp(t\boldsymbol{A})}{h}\\ &= \exp(t\boldsymbol{A})\lim_{h\to 0}\frac{\exp(h\boldsymbol{A}) - \boldsymbol{I}}{h}\\ &= \exp(t\boldsymbol{A})\boldsymbol{A};\end{aligned}$$

其中最后一个极限等于 $\boldsymbol{A}$ 是从 $\exp(h\boldsymbol{A})$ 的级数定义中得到的. 由于 $\boldsymbol{A}$ 和 $\exp(t\boldsymbol{A})$ 的级数中的每一项都交换, 从而与 $\exp(t\boldsymbol{A})$ 也交换, 这就证明了命题. $\blacksquare$

现在我们回到求解微分方程系统. 下面的结论可以看成是具有常系数的线性微分方程组的基本定理.

定理 设 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵. 则初值问题 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}, \boldsymbol{X}(0) = \boldsymbol{X}_0$ 的解为 $\boldsymbol{X}(t) = \exp(t\boldsymbol{A})\boldsymbol{X}_0$, 而且它是唯一解.

证明 上一个命题表明

$$\frac{\mathrm{d}}{\mathrm{d}t}(\exp(t\boldsymbol{A})\boldsymbol{X}_0) = \left(\frac{\mathrm{d}}{\mathrm{d}t}\exp(t\boldsymbol{A})\right)\boldsymbol{X}_0 = \boldsymbol{A}\exp(t\boldsymbol{A})\boldsymbol{X}_0.$$

并且, 由于 $\exp(0\boldsymbol{A})\boldsymbol{X}_0=\boldsymbol{X}_0$, 因而 $\exp(t\boldsymbol{A})\boldsymbol{X}_0$ 就是初值问题的一个解. 为了说明没有其他的解, 设 $\boldsymbol{Y}(t)$ 为另一个满足 $\boldsymbol{Y}(0)=\boldsymbol{X}_0$ 的解, 令

$$\boldsymbol{Z}(t)=\exp(-t\boldsymbol{A})\boldsymbol{Y}(t).$$

于是,

$$\begin{aligned}\boldsymbol{Z}'(t) &= \left(\frac{\mathrm{d}}{\mathrm{d}t}\exp(-t\boldsymbol{A})\right)\boldsymbol{Y}(t)+\exp(-t\boldsymbol{A})\boldsymbol{Y}'(t)\\ &= -\boldsymbol{A}\exp(-t\boldsymbol{A})\boldsymbol{Y}(t)+\exp(-t\boldsymbol{A})\boldsymbol{A}\boldsymbol{Y}(t)\\ &= \exp(-t\boldsymbol{A})(-\boldsymbol{A}+\boldsymbol{A})\boldsymbol{Y}(t)\\ &\equiv 0.\end{aligned}$$

从而 $\boldsymbol{Z}(t)$ 为一个常数, 再令 $t=0$ 可得 $\boldsymbol{Z}(t)=\boldsymbol{X}_0$, 于是就有 $\boldsymbol{Y}(t)=\exp(t\boldsymbol{A})\boldsymbol{X}_0$. 这就完成了定理的证明. ■

注意, 这个证明与 1.1 节中给出的证明是一样的, 唯一的区别就是字母 $\boldsymbol{A}$ 的含义变了.

例　考虑系统

$$\boldsymbol{X}'=\begin{pmatrix}\lambda & 1\\ 0 & \lambda\end{pmatrix}\boldsymbol{X}.$$

根据上面的定理, 它的通解为

$$\boldsymbol{X}(t)=\exp(t\boldsymbol{A})\boldsymbol{X}_0=\exp\begin{pmatrix}t\lambda & t\\ 0 & t\lambda\end{pmatrix}\boldsymbol{X}_0.$$

这其中的矩阵指数在前面已经算过, 于是我们有

$$\boldsymbol{X}(t)=\begin{pmatrix}\mathrm{e}^{t\lambda} & t\mathrm{e}^{t\lambda}\\ 0 & \mathrm{e}^{t\lambda}\end{pmatrix}\boldsymbol{X}_0.$$

可以看出, 这与第 3 章中的计算结果相吻合. ■

6.5　非自治线性系统

直到现在, 几乎所有碰到的线性微分方程系统都是自治的. 但是, 在应用中还是经常会出现一些形式的非自治系统. 其中之一就是如下形式的系统

$$\boldsymbol{X}'=\boldsymbol{A}(t)\boldsymbol{X},$$

其中 $\boldsymbol{A}(t)=[a_{ij}(t)]$ 为 $n\times n$ 矩阵, 它连续地依赖于时间. 当在以后章节中遇到变分方程时, 我们会再深入研究这类系统.

现在我们只关心另一个不同的非自治线性系统:

$$\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{G}(t),$$

其中 $\boldsymbol{A}$ 为常值的 $n\times n$ 矩阵, $\boldsymbol{G}:\mathbb{R}\to\mathbb{R}^n$ 显示地依赖于 t, 我们称之为**强迫项**. 这是一个一阶线性非自治方程系统.

例 (**受迫调和振子**) 如果调和振子系统中有一个外力作用, 决定运动的微分方程就变成

$$x''+bx'+kx=f(t),$$

其中 $f(t)$ 为外力的大小. 一个重要的特殊情形是其中的外力是时间的周期函数, 这可对应于, 比方说, 将质点 - 弹簧装置所在的桌子前后周期地移动. 作为一个系统, 受迫调和振子方程变成

$$\boldsymbol{X}'=\begin{pmatrix}0 & 1\\ -k & -b\end{pmatrix}\boldsymbol{X}+\boldsymbol{G}(t),\quad \text{其中}\quad \boldsymbol{G}(t)=\begin{pmatrix}0\\ f(t)\end{pmatrix}.$$

■

对一个非齐次的系统, 如果将其中的时间项扔掉后, 所得的系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 称为它的**齐次方程**. 我们已经知道如何找到这类系统的通解. 利用上节中的记号, 齐次系统满足初值条件 $\boldsymbol{X}(0)=\boldsymbol{X}_0$ 的解为

$$\boldsymbol{X}(t)=\exp(t\boldsymbol{A})\boldsymbol{X}_0,$$

而这也是齐次方程的通解.

为了找到非齐次方程的通解, 假设我们已经有了该方程的一个特解 $\boldsymbol{Z}(t)$. 于是, $\boldsymbol{Z}'(t)=\boldsymbol{A}\boldsymbol{Z}(t)+\boldsymbol{G}(t)$. 当 $\boldsymbol{X}(t)$ 为齐次方程的任一个解时, 则函数 $\boldsymbol{Y}(t)=\boldsymbol{X}(t)+\boldsymbol{Z}(t)$ 就是非齐次方程的另一个解. 这点可由如下的计算得出:

$$\begin{aligned}\boldsymbol{Y}'=\boldsymbol{X}'+\boldsymbol{Z}'&=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{A}\boldsymbol{Z}+\boldsymbol{G}(t)\\ &=\boldsymbol{A}(\boldsymbol{X}+\boldsymbol{Z})+\boldsymbol{G}(t)\\ &=\boldsymbol{A}\boldsymbol{Y}+\boldsymbol{G}(t).\end{aligned}$$

因而, 由于已经知道了齐次方程的所有解, 一旦找到了非齐次方程的哪怕只是一个特解, 我们就马上可以找到非齐次方程的通解. 通常, 我们可以通过猜解得到这样一个解 (微积分中, 这种方法通常称为**待定系数法**). 然而, 猜解并不总是可行. 下面的**参数变易法**却总是可行的, 当然, 这并不能保证我们总能计算其中的积分.

定理 (**参数变易法**) 考虑非齐次方程

$$\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{G}(t),$$

其中 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵, $\boldsymbol{G}(t)$ 为 t 的连续函数. 则

$$\boldsymbol{X}(t) = \exp(t\boldsymbol{A})\left(\boldsymbol{X}_0 + \int_0^t \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s\right)$$

是该方程满足 $\boldsymbol{X}(0) = \boldsymbol{X}_0$ 的一个解.

证明 对 $\boldsymbol{X}(t)$ 求导, 我们就有

$$\begin{aligned}\boldsymbol{X}'(t) = &\ \boldsymbol{A}\exp(t\boldsymbol{A})\left(\boldsymbol{X}_0 + \int_0^t \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s\right) + \exp(t\boldsymbol{A})\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s \\ = &\ \boldsymbol{A}\exp(t\boldsymbol{A})\left(\boldsymbol{X}_0 + \int_0^t \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s\right) + \boldsymbol{G}(t) \\ = &\ \boldsymbol{A}\boldsymbol{X}(t) + \boldsymbol{G}(t).\end{aligned}$$

■

下面我们来给出这个结果在周期受迫调和振子中的几个应用. 首先, 假设有一个外力为 $\cos t$ 的阻尼振子, 此时外力项的周期为 2π. 这个系统为

$$\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{G}(t),$$

其中 $\boldsymbol{G}(t) = (0, \cos t)$, $\boldsymbol{A}$ 为矩阵

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -k & -b \end{pmatrix},$$

其中 $b, k > 0$. 我们断言：这个系统有唯一一个以 2π 为周期的周期解. 为了证明这个断言, 我们首先需要找到一个满足条件 $\boldsymbol{X}(0) = \boldsymbol{X}_0 = \boldsymbol{X}(2\pi)$ 的解. 由参数变易法, 我们需要找到 $\boldsymbol{X}_0$ 使得

$$\boldsymbol{X}_0 = \exp(2\pi\boldsymbol{A})\boldsymbol{X}_0 + \exp(2\pi\boldsymbol{A})\int_0^{2\pi} \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s.$$

现在其中一项

$$\exp(2\pi\boldsymbol{A})\int_0^{2\pi} \exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s$$

为常值向量, 将它记为 $\boldsymbol{W}$. 这样我们就要求解方程

$$(\exp(2\pi\boldsymbol{A})-\boldsymbol{I})\boldsymbol{X}_0=-\boldsymbol{W}.$$

由于矩阵 $\exp(2\pi\boldsymbol{A})-\boldsymbol{I}$ 可逆, 这个方程有唯一解. 为证明这个矩阵的可逆性, 假设它是不可逆的, 则有非零向量 $\boldsymbol{V}$ 使得

$$(\exp(2\pi\boldsymbol{A})-\boldsymbol{I})\boldsymbol{V}=\boldsymbol{0},$$

或者, 换句话说, 矩阵 $\exp(2\pi\boldsymbol{A})$ 有一个特征值为 1. 但是, 由上节可知, $\exp(2\pi\boldsymbol{A})$ 的特征值都由 $\exp(2\pi\lambda_j)$ 给出, 其中 λ_j 为 $\boldsymbol{A}$ 的特征值. 由于每个 λ_j 的实部都小于 0, 从而 $\exp(2\pi\lambda_j)$ 小于 1, 于是矩阵 $\exp(2\pi\boldsymbol{A})-\boldsymbol{I}$ 的确是可逆的. 这样导致 2π 周期解的唯一初值就是

$$\boldsymbol{X}_0=(\exp(2\pi\boldsymbol{A})-\boldsymbol{I})^{-1}(-\boldsymbol{W}).$$

令 $\boldsymbol{X}(t)$ 为满足 $\boldsymbol{X}(0)=\boldsymbol{X}_0$ 的周期解. 这个解称为**定态解**. 如果 $\boldsymbol{Y}_0$ 是其他的初值条件, 我们可将它写成 $\boldsymbol{Y}_0=(\boldsymbol{Y}_0-\boldsymbol{X}_0)+\boldsymbol{X}_0$, 从而过 $\boldsymbol{Y}_0$ 的解为

$$\begin{aligned}\boldsymbol{Y}(t)&=\exp(t\boldsymbol{A})(\boldsymbol{Y}_0-\boldsymbol{X}_0)+\exp(t\boldsymbol{A})\boldsymbol{X}_0+\exp(t\boldsymbol{A})\int_0^t\exp(-s\boldsymbol{A})\boldsymbol{G}(s)\mathrm{d}s\\&=\exp(t\boldsymbol{A})(\boldsymbol{Y}_0-\boldsymbol{X}_0)+\boldsymbol{X}(t).\end{aligned}$$

对于表达式中的第一项, 由于它是齐次方程的一个解, 因而在 $t\to\infty$ 时, 趋于 0. 于是, 这个系统的每个解在 $t\to\infty$ 时都趋于定态解. 在物理上, 这一点是很明显的: 带阻尼的 (非受迫) 调和振子的运动趋于平衡态, 剩下的运动就是由周期外力引起的. 这样, 我们就证明了:

定理 考虑受迫带阻尼的调和振子方程

$$x''+bx'+kx=\cos t,$$

其中 $k,b>0$. 则这个方程的所有解都趋于一个以 2π 为周期的定态解. ■

现在考虑一个受迫的无阻尼调和振子的特例:

$$\boldsymbol{X}'=\begin{pmatrix}0&1\\-1&0\end{pmatrix}\boldsymbol{X}+\begin{pmatrix}0\\\cos\omega t\end{pmatrix},$$

其中, 现在外力的周期为 $2\pi/\omega,\omega\neq\pm1$. 记

$$\boldsymbol{A}=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

齐次方程的解为

$$\boldsymbol{X}(t)=\exp(t\boldsymbol{A})\boldsymbol{X}_0=\begin{pmatrix}\cos t & \sin t\\ -\sin t & \cos t\end{pmatrix}\boldsymbol{X}_0.$$

由参数变易法可知，非齐次方程从原点出发的解为

$$\begin{aligned}\boldsymbol{Y}(t)&=\exp(t\boldsymbol{A})\int_0^t\exp(-s\boldsymbol{A})\begin{pmatrix}0\\ \cos\omega s\end{pmatrix}\mathrm{d}s\\ &=\exp(t\boldsymbol{A})\int_0^t\begin{pmatrix}\cos s & -\sin s\\ \sin s & \cos s\end{pmatrix}\begin{pmatrix}0\\ \cos\omega s\end{pmatrix}\mathrm{d}s\\ &=\exp(t\boldsymbol{A})\int_0^t\begin{pmatrix}-\sin s\cos\omega s\\ \cos s\cos\omega s\end{pmatrix}\mathrm{d}s\\ &=\frac{1}{2}\exp(t\boldsymbol{A})\int_0^t\begin{pmatrix}\sin(\omega-1)s-\sin(\omega+1)s\\ \cos(\omega-1)s+\cos(\omega+1)s\end{pmatrix}\mathrm{d}s.\end{aligned}$$

回想到

$$\exp(t\boldsymbol{A})=\begin{pmatrix}\cos t & \sin t\\ -\sin t & \cos t\end{pmatrix},$$

然后再利用 $\omega\neq\pm1$ 的事实，算出其中的积分值再经过一个长的计算过程后就可得到

$$\begin{aligned}\boldsymbol{Y}(t)&=\frac{1}{2}\exp(t\boldsymbol{A})\begin{pmatrix}\dfrac{-\cos(\omega-1)t}{\omega-1}+\dfrac{\cos(\omega+1)t}{\omega+1}\\ \dfrac{\sin(\omega-1)t}{\omega-1}+\dfrac{\sin(\omega+1)t}{\omega+1}\end{pmatrix}+\exp(t\boldsymbol{A})\begin{pmatrix}(\omega^2-1)^{-1}\\ 0\end{pmatrix}\\ &=\frac{1}{\omega^2-1}\begin{pmatrix}-\cos\omega t\\ \omega\sin\omega t\end{pmatrix}+\exp(t\boldsymbol{A})\begin{pmatrix}(\omega^2-1)^{-1}\\ 0\end{pmatrix}.\end{aligned}$$

这样就得到原方程的通解为

$$\boldsymbol{Y}(t)=\exp(t\boldsymbol{A})\left(\boldsymbol{X}_0+\begin{pmatrix}(\omega^2-1)^{-1}\\ 0\end{pmatrix}\right)+\frac{1}{\omega^2-1}\begin{pmatrix}-\cos\omega t\\ \omega\sin\omega t\end{pmatrix}.$$

这个表达式中的第一项是以 2π 为周期的，而第二项则是以 $2\pi/\omega$ 为周期的. 与有阻尼的情形不同，现在的解不一定会有周期运动. 事实上，这个解为周期的当且仅当 ω 是一个有理数. 当 ω 为无理数时，这个运动是拟周期的，这正好与 6.2 节中看到的情形是一样的.

习　题

1. 找出系统 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的通解, 其中 $\boldsymbol{A}$ 为

(a) $\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$　(b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$　(c) $\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$　(e) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$　(f) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

(g) $\begin{pmatrix} 1 & 0 & -1 \\ -1 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$　(h) $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

2. 考虑线性系统

$$\boldsymbol{X}' = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \boldsymbol{X},$$

其中 $\lambda_3 < \lambda_2 < \lambda_1 < 0$. 描述过任一初值的解是如何趋于原点的.

3. 举例说明存在一个 3×3 的矩阵 $\boldsymbol{A}$ 使得 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 的所有非平衡解都是以 2π 为周期的. 简略地画出它的相图.

4. 找出

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \boldsymbol{X}$$

的通解.

5. 考虑依赖于两个参数 a, b 的系统

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ a & 0 & 0 \end{pmatrix} \boldsymbol{X}.$$

(a) 找出这个系统的通解.

(b) 在 ab 平面上作出具有不同类型相图的系统所对应的区域.

6. 考虑依赖于两个参数 a, b 的系统

$$\boldsymbol{X}' = \begin{pmatrix} a & 0 & b \\ 0 & b & 0 \\ -b & 0 & a \end{pmatrix} \boldsymbol{X}.$$

(a) 找出这个系统的通解.

(b) 在 ab 平面上作出具有不同类型相图的系统所对应的区域.

7. **耦合调和振子** 在这个系列的习题中要求你将 6.2 节中关于调和振子的内容推广到耦合振子情形. 假设两个质点 m_1, m_2 被弹簧以及墙体连接起来, 如图 6.10. 将 m_j 连接到墙体的弹簧的弹性系数都是 k_1, 而连接 m_1 和 m_2 的弹簧的弹性系数为 k_2. 这里耦合意味着每个质点的运动都要对另一个质点的行为产生影响.

图 6.10　一个耦合振子

用 x_j 记质点离开它的平衡位置的位移, 并且假设两个质点的质量都为 1. 则耦合振子的微分方程为

$$x_1'' = -(k_1 + k_2)x_1 + k_2 x_2$$

$$x_2'' = k_2 x_1 - (k_1 + k_2)x_2.$$

这两个方程可以如下得到. 如果 m_1 向右边移动 $(x_1 > 0)$, 则左边的弹簧被拉长, 从而对 m_1 产生一个大小为 $-k_1 x_1$ 的回复力, 同时中间的弹簧被压缩, 从而对 m_1 产生一个大小为 $-k_2 x_1$ 的回复力. 如果右边的弹簧被拉长, 则中间的弹簧被压缩, 从而对 m_1 产生一个大小为 $k_2 x_2$ 的回复力 (因为 $x_2 < 0$). 在 m_2 上的作用力可以类似得到.

(a) 将这两个方程写成一个一阶线性系统.

(b) 确定相应矩阵的特征值和特征向量.

(c) 找出通解.

(d) 记 $\omega_1 = \sqrt{k_1}, \omega_2 = \sqrt{k_1 + 2k_2}$. 解的周期性和 ω_j 之间有何关系? 请给出证明.

8. 假设 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 其中 $\boldsymbol{A}$ 为一 4×4 矩阵, 特征值为 $\pm i\sqrt{2}$ 和 $\pm i\sqrt{3}$. 描述这个流.

9. 假设 $\boldsymbol{X}' = \boldsymbol{AX}$, 其中 $\boldsymbol{A}$ 为一 4×4 矩阵, 特征值为 $\pm\mathrm{i}$ 和 $-1\pm\mathrm{i}$. 描述这个流.
10. 假设 $\boldsymbol{X}' = \boldsymbol{AX}$, 其中 $\boldsymbol{A}$ 为一 4×4 矩阵, 特征值为 $\pm\mathrm{i}$ 和 ±1. 描述这个流.
11. 考虑系统 $\boldsymbol{X}' = \boldsymbol{AX}$, 其中 $\boldsymbol{X} = (x_1,\cdots,x_6)$,

$$\boldsymbol{A} = \begin{pmatrix} 0 & \omega_1 & & & & \\ -\omega_1 & 0 & & & & \\ & & 0 & \omega_2 & & \\ & & -\omega_2 & 0 & & \\ & & & & -1 & \\ & & & & & 1 \end{pmatrix},$$

并且 ω_1/ω_2 为无理数. 在下面几种情形描述解的定性行为, 在时刻 0, 除了下面提到的外, 其他 x_j 都非零.

(a) $x_6 = 0$; (b) $x_5 = 0$; (c) $x_3 = x_4 = x_5 = 0$; (d) $x_3 = x_4 = x_5 = x_6 = 0$.

12. 计算下列矩阵的指数：

(a) $\begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$ (b) $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ (c) $\begin{pmatrix} 2 & -1 \\ 0 & 2 \end{pmatrix}$ (d) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$

(e) $\begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ (f) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 1 & 3 \end{pmatrix}$ (g) $\begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix}$

(h) $\begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}$ (i) $\begin{pmatrix} 1+\mathrm{i} & 0 \\ 2 & 1+\mathrm{i} \end{pmatrix}$ (j) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

13. 举例说明存在两个矩阵 $\boldsymbol{A}, \boldsymbol{B}$ 使得

$$\exp(\boldsymbol{A}+\boldsymbol{B}) \neq \exp(\boldsymbol{A})\exp(\boldsymbol{B}).$$

14. 证明：如果 $\boldsymbol{AB} = \boldsymbol{BA}$, 则有

(a) $\exp(\boldsymbol{A})\exp(\boldsymbol{B}) = \exp(\boldsymbol{B})\exp(\boldsymbol{A})$; (b) $\exp(\boldsymbol{A})\boldsymbol{B} = \boldsymbol{B}\exp(\boldsymbol{A})$.

15. 考虑三个调和振子

$$x_1'' = -x_1$$

$$x_2'' = -2x_2$$

$$x_3'' = -\omega^2 x_3$$

其中 ω 为无理数. 试描述这个六维系统的解的定性行为.

第 7 章　非线性系统

在本章, 我们开始研究非线性微分方程. 对于线性 (常系数) 系统, 我们总可以找到任一初值问题的显式解, 对于非线性系统, 这种情况是相当少见的. 事实上, 一些基本性质, 例如解的存在唯一性, 在线性情形是显然的, 而在非线性情形则不再成立. 我们将看到, 有些非线性系统甚至对任何初值问题都无解. 另一方面, 有些系统却有无穷多个不同的解. 而且即使找到了这样系统的一个解, 它也不一定对所有时刻都有定义. 例如, 解有可能在有限时间就趋于 ∞. 随之而来, 问题就产生了：例如, 当系统的初始条件发生哪怕是一点微小变化时, 将会发生什么? 相应的解会连续变化吗? 所有的这些问题对线性系统都是清楚的, 而对非线性系统则未必. 这意味着非线性微分方程的基本理论要比线性系统复杂得多.

实际中产生的大部分的非线性系统都是"好的", 也就是说, 它们具有解的存在唯一性, 当初始条件改变时, 解也会随之而连续变化, 其他的"自然"性质也将发生变化. 这样, 我们就得面临一个选择：给定一个非线性系统, 我们要么简单地向前进行, 寄希望 (或者去验证) 在每一种特定情形, 系统解的行为都很好；要么, 在此刻停下来, 花很多时间来阐述一些必要的假设, 以保证给定的非线性系统解的行为很好.

在本书中, 我们采取一种综合的路线来进行. 在本章, 我们将精确陈述很多理论性的结果, 以保证微分方程解的行为很好. 我们也将给出一些例子来说明在什么情况下, 这些结果是如何失效的. 但是, 在此我们不打算证明这些定理, 而是将所有的技术性证明留到第 17 章来完成, 这主要是因为理解这些内容需要大量坚实的实分析原理背景. 在以后的章节中, 我们要用到此处陈述的结果, 但如果读者主要是对微分方程的应用或对如何分析一些特定非线性系统感兴趣的话, 则大可不必陷入到这些细节中. 而对技术细节感兴趣的读者可以现在就去看看第 17 章的内容.

7.1　动力系统

前面已经提到, 大多数的非线性微分方程系统不可能解析地求解. 原因之一就是, 我们没有足够多具有特定名称的函数来写出这些系统的显式解. 同样困难的是, 我们将看到, 高维系统可能产生混沌行为, 这种性质使得即使知道了显式解, 也不会对我们理解系统的整体行为有什么实质性的帮助. 于是, 为了理解这些系统, 我们被迫去寻找另外的办法, 这就是在动力系统领域出现的各种技巧. 我们将

综合应用分析的、几何的以及拓扑的技巧来得到有关这些方程解的行为的严格结果.

我们先罗列一些与动力系统有关的术语, 在前面几章, 我们已经在各种不同的场合引入过这些术语. **动力系统**是一种描述一个给定空间 $\mathcal{S}$ 中所有点随时间旅程的方法. 空间 $\mathcal{S}$ 可以看成是, 比如, 某个物理系统的状态空间. 数学上, $\mathcal{S}$ 可能是欧几里得空间, 也可能是欧几里得空间的一个开子集, 或一些其他空间, 如 $\mathbb{R}^3$ 中的曲面. 当我们考虑力学中产生的动力系统时, 空间 $\mathcal{S}$ 将是系统所有可能的位置和速度. 为了简单起见, 我们将始终假设空间 $\mathcal{S}$ 就是欧几里得空间 $\mathbb{R}^n$, 但在某些情形, 重要的动力行为将只局限在 $\mathbb{R}^n$ 的一个特定子集上.

给定一个初始位置 $\boldsymbol{X} \in \mathbb{R}^n$, $\mathbb{R}^n$ 上的一个动力系统会告诉我们 $\boldsymbol{X}$ 在 1 个单位时间后在哪里, 2 个单位时间在哪里, 等等. 我们将 $\boldsymbol{X}$ 的这些新位置记为 $\boldsymbol{X}_1, \boldsymbol{X}_2$, 等等. 在零时刻, $\boldsymbol{X}$ 的位置为 $\boldsymbol{X}_0$. 在零时刻前一个单位时刻, $\boldsymbol{X}$ 的位置为 $\boldsymbol{X}_{-1}$. 一般地, $\boldsymbol{X}$ 的"轨线"由 $\boldsymbol{X}_t$ 给出. 如果只是在整数时间值测量位置 $\boldsymbol{X}_t$, 我们就得到一个**离散动力系统**的例子, 我们将在第 15 章研究离散动力系统. 如果在连续时间 $t \in \mathbb{R}$ 测量位置 $\boldsymbol{X}_t$, 我们就得到一个**连续动力系统**. 如果系统以一种连续可微的方式依赖于时间, 我们就得到一个**光滑动力系统**. 这是在微分方程系统研究中产生的三种主要类型的动力系统, 它们将构成第 8 章到第 14 章的主干.

将 t 带到 $\boldsymbol{X}_t$ 的函数要么是得到 $\mathbb{R}^n$ 中的一列点, 要么是得到 $\mathbb{R}^n$ 中的一条曲线, 无论哪种情形, 它都纪录了当时间从 $-\infty$ 跑到 ∞ 时, $\boldsymbol{X}$ 的生活历史. 动力系统的不同分支需要对 $\boldsymbol{X}_t$ 怎样依赖于 t 作不同的假设. 例如, 遍历理论需要假设这样的函数保持 $\mathbb{R}^n$ 上的一个测度不变; 拓扑动力系统则只须假设 $\boldsymbol{X}_t$ 是连续变化的. 在微分方程情形, 我们通常假设 $\boldsymbol{X}_t$ 是连续可微的. 对每个时刻 t, 我们都可以定义映射 $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$, 它将 $\boldsymbol{X}$ 映到 $\boldsymbol{X}_t$. 由于我们可将 $\boldsymbol{X}_t$ 理解成随时间变化的一个状态, 期望 ϕ_t 是以 ϕ_{-t} 为逆就是很自然的事. 同样地, ϕ_0 必定就是恒等函数 $\phi_0(\boldsymbol{X}) = \boldsymbol{X}$, 而 $\phi_t(\phi_s(\boldsymbol{X})) = \phi_{t+s}(\boldsymbol{X})$ 也是一个自然的条件. 我们将所有的这些都正式地陈述在下面的定义中.

定义 $\mathbb{R}^n$ 上的一个**光滑动力系统**是指这样的一个连续可微函数 $\phi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, 其中 $\phi(t, \boldsymbol{X}) = \phi_t(\boldsymbol{X})$ 满足下面的两条:

(1) $\phi_0 : \mathbb{R}^n \to \mathbb{R}^n$ 为恒等函数: $\phi_0(\boldsymbol{X}_0) = \boldsymbol{X}_0$;

(2) 对于所有的 $t, s \in \mathbb{R}$, 复合函数 $\phi_t \circ \phi_s = \phi_{t+s}$. ■

回忆一下, 一个函数是连续可微的, 如果在整个定义域上, 它的所有偏导数都存在并且连续. 习惯上, 一个连续可微函数是指一个 C^1 函数; 如果函数是 k 次连续可微的, 它就称为一个 C^k 函数. 注意, 上面的定义蕴涵着, 对于每一个 t, $\phi_t : \mathbb{R}^n \to \mathbb{R}^n$ 都是 C^1 的, 而且有 C^1 的逆 ϕ_{-t}(在第 2 条中令 $s = -t$ 即可).

例 对于一阶微分方程 $x' = ax$, 函数 $\phi_t(x_0) = x_0 \exp(at)$ 给出了这个方程的所有

解, 而且也定义了 $\mathbb{R}$ 上的一个光滑动力系统. ■

例 令 $\boldsymbol{A}$ 为一 $n\times n$ 矩阵. 则函数 $\phi_t(\boldsymbol{X}_0)=\exp(t\boldsymbol{A})\boldsymbol{X}_0$ 就定义了 $\mathbb{R}^n$ 上的一个光滑动力系统. 显然, $\phi_0=\exp(0)=\boldsymbol{I}$, 而且在上一章中已经看到:

$$\phi_{t+s}=\exp((t+s)\boldsymbol{A})=(\exp(t\boldsymbol{A}))(\exp(s\boldsymbol{A}))=\phi_t\circ\phi_s. \quad ■$$

可以看出, 上面两个例子都是与微分方程系统 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 紧密相关的. 一般地, 从每个光滑动力系统都可以如下得到 $\mathbb{R}^n$ 上的一个向量场: 给定 ϕ_t, 令

$$\boldsymbol{F}(\boldsymbol{X})=\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0}\phi_t(\boldsymbol{X}).$$

于是, ϕ_t 恰好就是与 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的流相应的时间 t 映射.

反过来, 如果微分方程 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的流对所有的时间, 其时间 t 映射都有良好的定义, 而且是连续可微的, 则这个微分方程就生成了一个光滑动力系统. 很遗憾, 事情并不总是如此.

7.2 存在唯一性定理

我们现在转到微分方程的基本定理 —— 存在唯一性定理. 考虑微分方程系统

$$\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X}),$$

其中 $\boldsymbol{F}:\mathbb{R}^n\to\mathbb{R}^n$. 回忆一下, 此系统的一个解是指定义在某一区间 $J\subset\mathbb{R}$ 上的函数 $\boldsymbol{X}:J\to\mathbb{R}^n$, 使得对所有的 $t\in J$, 都有

$$\boldsymbol{X}'(t)=\boldsymbol{F}(\boldsymbol{X}(t)).$$

几何上, $\boldsymbol{X}(t)$ 是 $\mathbb{R}^n$ 中的一条曲线, 对于所有的 $t\in J$, 切向量 $\boldsymbol{X}'(t)$ 都存在并且等于 $\boldsymbol{F}(\boldsymbol{X}(t))$. 像前面几章一样, 我们认为这个向量的基点位于 $\boldsymbol{X}(t)$, 于是映射 $\boldsymbol{F}:\mathbb{R}^n\to\mathbb{R}^n$ 就定义了 $\mathbb{R}^n$ 上的一个向量场. 解 $\boldsymbol{X}:J\to\mathbb{R}^n$ 的一个**初始条件**或一个**初值**是指形如 $\boldsymbol{X}(t_0)=\boldsymbol{X}_0$ 的一个规定, 其中 $t_0\in J,\boldsymbol{X}_0\in\mathbb{R}^n$. 为了简单, 我们通常取 $t_0=0$. 微分方程的一个主要问题就是寻找任一**初值问题**的解, 即对每个 $\boldsymbol{X}_0\in\mathbb{R}^n$, 确定系统的解使得它满足初始条件 $\boldsymbol{X}(0)=\boldsymbol{X}_0$.

很遗憾, 非线性微分方程对某些初始条件可能会没有解.

例 考虑下面简单的一阶微分方程

$$x'=\begin{cases}1 & \text{如果 } x<0\\ -1 & \text{如果 } x\geqslant 0.\end{cases}$$

$\mathbb{R}$ 上的这个向量场在 $x \geqslant 0$ 时指向左边, 而在 $x < 0$ 时指向右边. 因此, 没有解满足初始条件 $x(0) = 0$. 事实上, 如果有这样的一个解, 由于 $x'(0) = -1$, 它在初始时一定是递减的, 但是, 对所有取负值的 x, 解又必须是递增的, 这不可能发生. 进一步, 还可注意到, 任何一个解都不可能对所有时间都有定义. 例如, 当 $x_0 > 0$ 时, 过 x_0 的解为 $x(t) = x_0 - t$, 但根据刚才同样的理由, 这个解只在 $-\infty < t < x_0$ 时有意义.

这个例子中的问题在于向量场在 0 处并不连续; 只要向量场在某点出现间断, 我们就可能面对相邻的向量却指向"相反"的方向, 这样就使得解在这些坏点处中断. ■

对于非线性微分方程, 除了解的存在性问题之外, 我们还会碰到, 某些方程对于同一个初值问题可能会有许多不同的解.

例 考虑微分方程

$$x' = 3x^{2/3}.$$

恒等于零的函数 $u : \mathbb{R} \to \mathbb{R}, u(t) \equiv 0$ 显然是满足初始条件 $u(0) = 0$ 的一个解, 而 $u_0(t) = t^3$ 也是满足这个初始条件的一个解. 进一步, 对任意的 $\tau > 0$, 由下式给出的函数

$$u_\tau(t) = \begin{cases} 0 & \text{如果 } t \leqslant \tau \\ (t-\tau)^3 & \text{如果 } t > \tau, \end{cases}$$

也是满足初始条件 $u_\tau(0) = 0$ 的解. 然而, 这个例子中的微分方程在 $x_0 = 0$ 处是连续的, 现在的问题在于 $3x^{2/3}$ 在这点处不可微. ■

从这两个例子中可以清楚地看到, 为了保证解的存在唯一性, 必须要对函数 $\boldsymbol{F}$ 加上某些条件. 在第一个例子中, $\boldsymbol{F}$ 在问题点 0 处不连续, 而在第二个例子中, $\boldsymbol{F}$ 在 0 处不可微. 这表明 $\boldsymbol{F}$ 连续可微的假设可能就足以保证解的存在且唯一, 这点我们随后就会看到. 幸运的是, 应用中产生的不连续可微的微分方程还是很少的, 因而, 对于给定初始条件, 不存在解或解不唯一的现象是相当例外的.

下面就是常微分方程的基本局部定理.

存在唯一性定理 考虑初值问题

$$\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X}),\ \boldsymbol{X}(t_0) = \boldsymbol{X}_0,$$

其中 $\boldsymbol{X}_0 \in \mathbb{R}^n$. 假设 $\boldsymbol{F} : \mathbb{R}^n \to \mathbb{R}^n$ 是 C^1 的, 则, 首先该初值问题有一个解, 其次, 只有一个这样的解. 更准确地说, 存在 $a > 0$, 以及该微分方程满足初始条件 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$ 的唯一解

$$\boldsymbol{X} : (t_0 - a, t_0 + a) \to \mathbb{R}^n.$$

■

该定理的重要证明见第 17 章.

在此, 如果不管细节的话, 定理的证明依赖于一个重要的技巧: **毕卡迭代**. 在继续讨论之前, 我们先来解释定理证明中毕卡迭代法在几个特例中是如何发挥作用的. 迭代过程背后的基本想法是构造趋于微分方程解的一个函数序列. 函数序列 $u_k(t)$ 是归纳地定义的, 首先令 $u_0(t) = x_0$, 这里的 x_0 就是给定的初始条件, 然后令

$$u_{k+1}(t) = x_0 + \int_0^t u_k(s)\mathrm{d}s.$$

例 考虑简单的微分方程 $x' = x$. 我们将造出这个方程满足 $x(0) = x_0$ 的解. 当然, 我们已经知道这个解就是由 $x(t) = x_0\mathrm{e}^t$ 给出的. 我们将构造一个函数序列 $u_k(t)$, 使得当 $k \to \infty$ 时, 序列 $u_k(t)$ 趋于实际的解 $x(t)$.

我们从

$$u_0(t) = x_0$$

开始, 其中 x_0 就是给定的初值. 然后, 我们令

$$u_1(t) = x_0 + \int_0^t u_0(s)\mathrm{d}s = x_0 + \int_0^t x_0\mathrm{d}s,$$

于是, $u_1(t) = x_0 + tx_0$. 给出 u_1 后, 我们接着定义

$$u_2(t) = x_0 + \int_0^t u_1(s)\mathrm{d}s = x_0 + \int_0^t (x_0 + sx_0)\mathrm{d}s,$$

于是, $u_2(t) = x_0 + tx_0 + \frac{t^2}{2}x_0$. 你或许已经知道这将走向哪里. 归纳地, 我们令

$$u_{k+1}(t) = x_0 + \int_0^t u_k(s)\mathrm{d}s,$$

于是就有

$$u_{k+1}(t) = x_0 \sum_{i=0}^{k+1} \frac{t^i}{i!}.$$

当 $k \to \infty$ 时, $u_k(t)$ 收敛到

$$x_0 \sum_{i=0}^{\infty} \frac{t^i}{i!} = x_0\mathrm{e}^t = x(t),$$

这正好就是我们原来方程的解. ■

例 我们来看一个毕卡迭代法应用到微分方程系统的例子. 考虑线性系统

$$\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X}) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \boldsymbol{X},$$

其初始条件为 $\boldsymbol{X}(0)=(1,0)$. 前面已经看到, 这个初值问题的解为

$$\boldsymbol{X}(t)=\begin{pmatrix}\cos t\\ -\sin t\end{pmatrix}.$$

运用毕卡迭代, 我们有

$$\begin{aligned}
U_0(t)&=\begin{pmatrix}1\\0\end{pmatrix}\\
U_1(t)&=\begin{pmatrix}1\\0\end{pmatrix}+\int_0^t F\begin{pmatrix}1\\0\end{pmatrix}\mathrm{d}s=\begin{pmatrix}1\\0\end{pmatrix}+\int_0^t\begin{pmatrix}0\\-1\end{pmatrix}\mathrm{d}s=\begin{pmatrix}1\\-t\end{pmatrix}\\
U_2(t)&=\begin{pmatrix}1\\0\end{pmatrix}+\int_0^t\begin{pmatrix}-s\\-1\end{pmatrix}\mathrm{d}s=\begin{pmatrix}1-t^2/2\\-t\end{pmatrix}\\
U_3(t)&=\begin{pmatrix}1\\0\end{pmatrix}+\int_0^t\begin{pmatrix}-s\\-1+s^2/2\end{pmatrix}\mathrm{d}s=\begin{pmatrix}1-t^2/2\\-t+t^3/3!\end{pmatrix}\\
U_4(t)&=\begin{pmatrix}1-t^2/2+t^4/4!\\-t+t^3/3!\end{pmatrix},
\end{aligned}$$

这样, 我们就看到了正弦函数和余弦函数的无穷级数在这个迭代中渐渐露出. ■

现在我们假设微分方程 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 有两个解 $\boldsymbol{Y}(t),\boldsymbol{Z}(t)$, 而且它们满足 $\boldsymbol{Y}(t_0)=\boldsymbol{Z}(t_0)$. 假设这两个解都是定义在区间 J 上. 存在唯一性定理保证了, 对于包含 t_0 的一个区间上的所有 t, 都有 $\boldsymbol{Y}(t)=\boldsymbol{Z}(t)$, 表面上, 这个区间可能会比 J 要来得小. 然而, 情况并非如此. 为了看出这一点, 假设 J^* 是满足 $\boldsymbol{Y}(t)=\boldsymbol{Z}(t)$ 的最大区间. 令 t_1 为 J^* 的一个端点. 根据连续性, 我们有 $\boldsymbol{Y}(t_1)=\boldsymbol{Z}(t_1)$. 则存在唯一性定理保证, $\boldsymbol{Y}(t)$ 和 $\boldsymbol{Z}(t)$ 实际上在一个包含 t_1 的开区间上都相等, 这与 J^* 是两个解相等的最大区间相矛盾.

因而我们总可以假设唯一解定义在一个最大的时间区间上. 然而这并不能保证 $\boldsymbol{X}(t)$ 对所有时刻都有定义, 不管 $\boldsymbol{F}(\boldsymbol{X})$ 是多么“好”.

例 考虑 $\mathbb{R}$ 上的微分方程

$$x'=1+x^2.$$

这个方程以函数 $x(t)=\tan(t+c)$ 作为它的解, 其中 c 为一常数. 由于当 $t\to-c\pm\pi/2$ 时, $x(t)\to\pm\infty$, 这个函数不能延拓到区间

$$-c-\frac{\pi}{2}<t<-c+\frac{\pi}{2}$$

之外. ■

这个例子很有代表性, 因为我们有

定理 假设 $U \subset \mathbb{R}^n$ 为一开集, 并假设 $\boldsymbol{F}: U \to \mathbb{R}^n$ 是 C^1 的. 令 $\boldsymbol{X}(t)$ 是 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的一个定义在最大开区间 $J = (\alpha, \beta) \subset \mathbb{R}$ 上的解, 其中 $\beta < \infty$. 则, 对任意的有界闭集 $K \subset U$, 总存在某个 $t \in (\alpha, \beta)$ 使得 $\boldsymbol{X}(t) \notin K$. ■

这个定理断言, 如果解 $\boldsymbol{X}(t)$ 不能延拓到一个更大的时间区间, 则这个解就会离开 U 中的任何有界闭集. 这意味着, 当 $t \to \beta$ 时, $\boldsymbol{X}(t)$ 可以任意接近 U 的边界. 同样的结果在 $t \to \alpha$ 时也成立.

7.3 解的连续依赖性

为了使得解的存在唯一性定理在各种物理 (甚至是数学) 意义下都有意义, 还需要加上解 $\boldsymbol{X}(t)$ 对初始条件 $\boldsymbol{X}(0)$ 的连续依赖性质. 下面的定理给出了这个性质的准确陈述.

定理 考虑微分方程 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, 其中 $\boldsymbol{F}: \mathbb{R}^n \to \mathbb{R}^n$ 为 C^1 的. 假设 $\boldsymbol{X}(t)$ 是该方程的一个解, 其在闭区间 $[t_0, t_1]$ 上有定义, 而且 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$. 则存在 $\boldsymbol{X}_0$ 的邻域 $U \subset \mathbb{R}^n$ 以及常数 K 使得, 如果 $\boldsymbol{Y}_0 \in U$, 则存在定义在 $[t_0, t_1]$ 上唯一的解 $\boldsymbol{Y}(t)$ 满足 $\boldsymbol{Y}(t_0) = \boldsymbol{Y}_0$. 而且, 对所有的 $t \in [t_0, t_1]$, $\boldsymbol{Y}(t)$ 满足

$$|\boldsymbol{Y}(t) - \boldsymbol{X}(t)| \leqslant K|\boldsymbol{Y}_0 - \boldsymbol{X}_0| \exp(K(t - t_0)).$$ ■

这个结果表明, 如果两个解 $\boldsymbol{X}(t)$ 和 $\boldsymbol{Y}(t)$ 出发时很接近, 则在 t 靠近 t_0 时, 它们一直很接近. 虽然这两个解可能分开, 但它们分开的速度不会超过指数增长速度. 特别地, 由于上面不等式的右端依赖于 $|\boldsymbol{Y}_0 - \boldsymbol{X}_0|$, 当假设它很小时, 我们就得到:

推论 **(对初始条件的连续依赖性)** 记 $\phi(t, \boldsymbol{X})$ 为系统 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的流, 其中 $\boldsymbol{F}$ 为 C^1 的. 则 ϕ 为 $\boldsymbol{X}$ 的连续函数. ■

例 令 $k > 0$, 我们知道系统

$$\boldsymbol{X}' = \begin{pmatrix} -1 & 0 \\ 0 & k \end{pmatrix} \boldsymbol{X}$$

满足 $\boldsymbol{X}(0) = (-1, 0)$ 的解为

$$\boldsymbol{X}(t) = (-\mathrm{e}^{-t}, 0).$$

对任一 $\eta \neq 0$, 记 $\boldsymbol{Y}_\eta(t)$ 为满足 $\boldsymbol{Y}_\eta(0) = (-1, \eta)$ 的解. 则,

$$\boldsymbol{Y}_\eta(t) = (-\mathrm{e}^{-t}, \eta \mathrm{e}^{kt}).$$

根据上面的定理, 我们有

$$|\boldsymbol{Y}_\eta(t) - \boldsymbol{X}(t)| = |\eta e^{kt} - 0| = |\eta - 0|e^{kt} = |\boldsymbol{Y}_\eta(0) - \boldsymbol{X}(0)|e^{kt}.$$

这些解 $\boldsymbol{Y}_\eta$ 的确会与 $\boldsymbol{X}(t)$ 分开 (见图 7.1), 但它们至多以指数速度分开. 而且, 对于任何固定的时间 t, 当 $\eta \to 0$ 时, $\boldsymbol{Y}_\eta(t) \to \boldsymbol{X}(t)$. ■

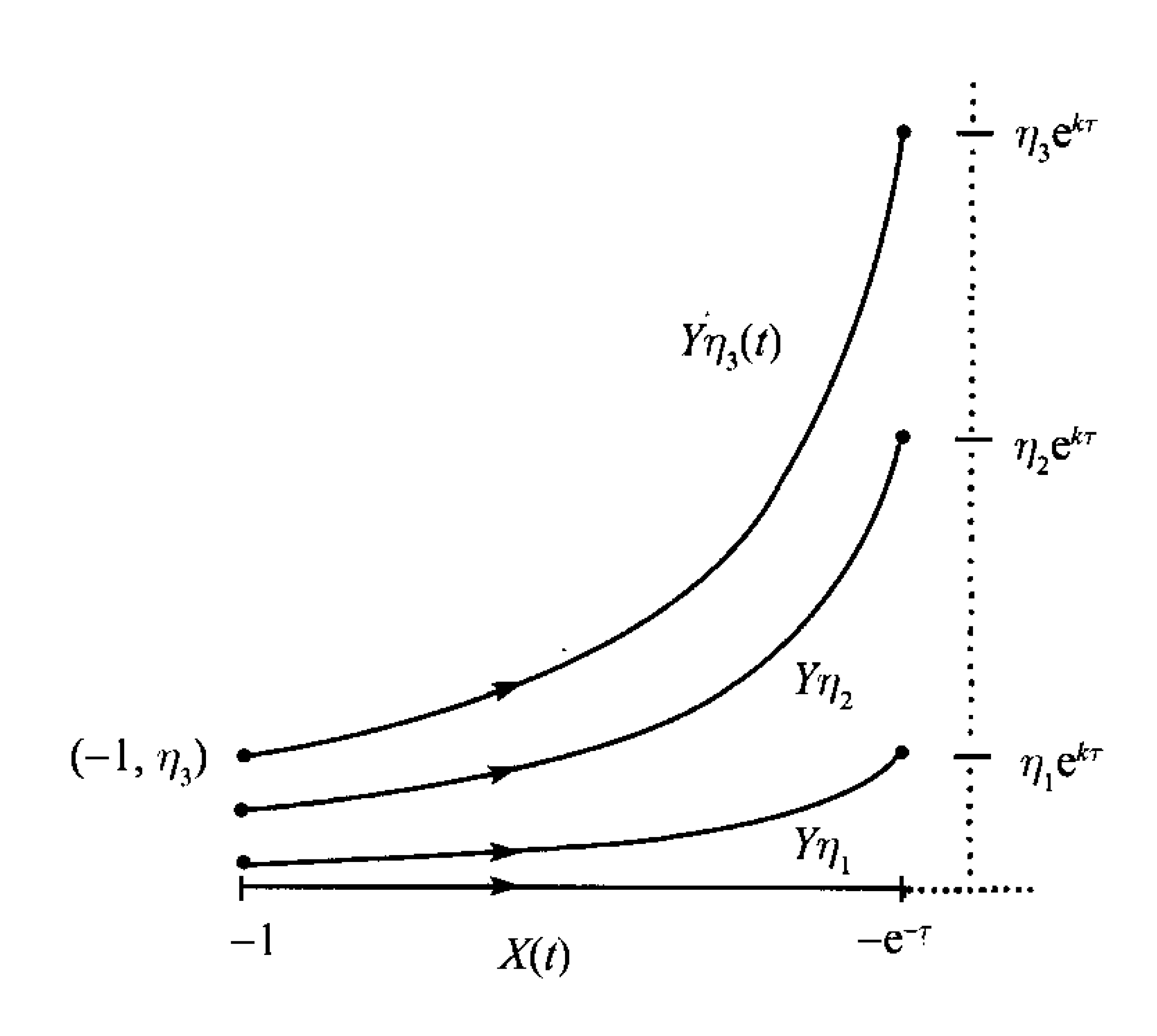


图 7.1 解 $\boldsymbol{Y}_\eta(t)$ 以指数速度与 $\boldsymbol{X}(t)$ 分开, 但仍对初始条件连续

微分方程常常依赖于参数. 例如, 调和振子方程依赖于参数 b(阻尼系数) 和 k(弹性系数). 那么, 一个自然的问题就是: 这些方程的解如何依赖于这些参数呢? 在前面的情形, 只要系统是以连续可微的方式依赖于这些参数, 则解就连续地依赖于这些参数. 利用一个特别的小技巧, 我们可以很容易地看出这一点. 假设系统

$$\boldsymbol{X}' = \boldsymbol{F}_a(\boldsymbol{X})$$

以 C^1 的方式依赖于参数 a. 我们来考虑一个 "人为的" 扩展的微分方程系统:

$$\begin{aligned}
x_1' &= f_1(x_1, \cdots, x_n, a) \\
&\ \vdots \\
x_n' &= f_n(x_1, \cdots, x_n, a) \\
a' &= 0.
\end{aligned}$$

这是 $n+1$ 个方程组的自治系统. 尽管这个系统的扩展部分看起来很平凡, 但如果对它运用前面的解对初始条件的连续依赖性结果, 可以得出原来系统的解对参数 a 也是连续依赖的.

定理　(**对参数的连续依赖性**)令 $\boldsymbol{X}' = \boldsymbol{F}_a(\boldsymbol{X})$ 为一个微分方程系统, 其中 $\boldsymbol{F}_a$ 对 $\boldsymbol{X}$ 和 a 都连续可微. 则这个系统的流也连续依赖于 a. ■

7.4 变分方程

考虑自治系统 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, 其中 $\boldsymbol{F}$ 像通常一样假设为 C^1 的. 该系统的流 $\phi(t, \boldsymbol{X})$ 同时为 t 和 $\boldsymbol{X}$ 的函数. 由上节的结果可知, ϕ 对变量 $\boldsymbol{X}$ 连续. 由于 $t \to \phi(t, \boldsymbol{X})$ 就是过 $\boldsymbol{X}$ 的解曲线, 我们还知道 ϕ 对变量 t 可微. 事实上, ϕ 对变量 $\boldsymbol{X}$ 也是可微的 (我们将在第 17 章给出其证明):

定理　(**流的可微性**) 考虑系统 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, 其中 $\boldsymbol{F}$ 为 C^1 的. 则, 该系统的流 $\phi(t, \boldsymbol{X})$ 为一 C^1 函数, 即 $\partial\phi/\partial t$ 和 $\partial\phi/\partial \boldsymbol{X}$ 都存在, 且对 t 和 $\boldsymbol{X}$ 连续. ■

注意, 一旦知道了过 $\boldsymbol{X}_0$ 的解, 则对任何 t, 我们总能够算出 $\partial\phi/\partial t$, 事实上, 我们有

$$\frac{\partial\phi}{\partial t}(t, \boldsymbol{X}_0) = \boldsymbol{F}(\phi(t, \boldsymbol{X}_0)).$$

另外, 我们还有

$$\frac{\partial\phi}{\partial \boldsymbol{X}}(t, \boldsymbol{X}_0) = \boldsymbol{D}\phi_t(\boldsymbol{X}_0),$$

其中 $\boldsymbol{D}\phi_t$ 为函数 $\boldsymbol{X} \to \phi_t(\boldsymbol{X})$ 的雅可比. 然而, 为了计算 $\partial\phi/\partial \boldsymbol{X}$, 除了要知道过 $\boldsymbol{X}_0$ 的解, 因为需要计算 ϕ_t 的各个分量的偏导数, 我们似乎还要知道所有经过附近初始位置的解. 但是, 通过引入沿过 $\boldsymbol{X}_0$ 的解的变分方程, 我们可以克服这个困难.

为了实现这一点, 需要先对非自治微分方程作简短讨论. 令 $\boldsymbol{A}(t)$ 为一族连续依赖于 t 的 $n \times n$ 矩阵. 系统

$$\boldsymbol{X}' = \boldsymbol{A}(t)\boldsymbol{X}$$

为一个线性非自治系统. 对这种类型的方程, 有如下的存在唯一性定理:

定理　令 $\boldsymbol{A}(t)$ 为定义在 $t \in [\alpha, \beta]$ 上的 $n \times n$ 矩阵的一个连续族. 则初值问题

$$\boldsymbol{X}' = \boldsymbol{A}(t)\boldsymbol{X}, \quad \boldsymbol{X}(t_0) = \boldsymbol{X}_0$$

具有唯一解, 而且解在整个区间 $[\alpha, \beta]$ 上都有定义. ■

注意, 这个定理有一些新的东西: 并不需要假设右端是 t 的 C^1 函数. $\boldsymbol{A}(t)$ 的连续性就足以保证解的存在唯一性.

例　考虑一阶线性非自治微分方程

$$x' = a(t)x.$$

利用第 1 章的方法, 很容易验证该方程满足 $x(0)=x_0$ 的唯一解为

$$x(t)=x_0\exp\left(\int_0^t a(s)\mathrm{d}s\right).$$

为了保证 $x'(t)=a(t)x(t)$, 我们只需要 $a(t)$ 为连续的, 而不需要它的可微性. ■

注意, 线性非自治微分方程的解满足线性叠加原理, 即, 如果 $\boldsymbol{Y}(t),\boldsymbol{Z}(t)$ 是这样一个系统的两个解, 则对任意的常数 α,β, $\alpha\boldsymbol{Y}(t)+\beta\boldsymbol{Z}(t)$ 也是它的解.

现在, 我们再回到自治非线性系统 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$. 令 $\boldsymbol{X}(t)$ 为该系统的一个特解, 假设在某个区间 $J=[\alpha,\beta]$ 上对 t 有定义. 固定 $t_0\in J$, 记 $\boldsymbol{X}(t_0)=\boldsymbol{X}_0$. 对每个 $t\in J$, 令

$$\boldsymbol{A}(t)=\boldsymbol{DF}_{\boldsymbol{X}(t)},$$

其中 $\boldsymbol{DF}_{\boldsymbol{X}(t)}$ 代表 $\boldsymbol{F}$ 在点 $\boldsymbol{X}(t)\in\mathbb{R}^n$ 处的雅可比矩阵. 由于 $\boldsymbol{F}$ 为 C^1 的, $\boldsymbol{A}(t)=\boldsymbol{DF}_{\boldsymbol{X}(t)}$ 是一族连续的 $n\times n$ 矩阵. 考虑非自治的线性方程

$$\boldsymbol{U}'=\boldsymbol{A}(t)\boldsymbol{U}.$$

该方程称为沿解 $\boldsymbol{X}(t)$ 的**变分方程**. 由上一个定理可知, 对所有的初始条件 $\boldsymbol{U}(t_0)=\boldsymbol{U}_0$, 该变分方程都有一个定义在整个 J 上的解.

这个方程的意义在于: 如果 $\boldsymbol{U}(t)$ 为变分方程满足 $\boldsymbol{U}(t_0)=\boldsymbol{U}_0$ 的解, 则只要 $\boldsymbol{U}_0$ 充分小, 函数

$$t\to\boldsymbol{X}(t)+\boldsymbol{U}(t)$$

就是原来自治方程满足初始条件 $\boldsymbol{Y}(t_0)=\boldsymbol{X}_0+\boldsymbol{U}_0$ 的解 $\boldsymbol{Y}(t)$ 的一个好的逼近. 这就是下面结果的内容.

命题 考虑系统 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$, 其中 $\boldsymbol{F}$ 为 C^1 的. 假设

(1) $\boldsymbol{X}(t)$ 为 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的一个解, 对所有的 $t\in[\alpha,\beta]$ 都有定义, 且 $\boldsymbol{X}(t_0)=\boldsymbol{X}_0$;

(2) $\boldsymbol{U}(t)$ 为沿 $\boldsymbol{X}(t)$ 的变分方程的解, 且 $\boldsymbol{U}(t_0)=\boldsymbol{U}_0$;

(3) $\boldsymbol{Y}(t)$ 为 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的解, 且 $\boldsymbol{Y}(t_0)=\boldsymbol{X}_0+\boldsymbol{U}_0$.

则

$$\lim_{\boldsymbol{U}_0\to 0}\frac{|\boldsymbol{Y}(t)-(\boldsymbol{X}(t)+\boldsymbol{U}(t))|}{|\boldsymbol{U}_0|}$$

对于 $t\in[\alpha,\beta]$ 一致收敛到 0. ■

严格来说就是, 对任意的 $\epsilon>0$, 存在 $\delta>0$ 使得当 $|\boldsymbol{U}_0|\leqslant\delta$ 时, 对所有的 $t\in[\alpha,\beta]$ 都有

$$|\boldsymbol{Y}(t)-(\boldsymbol{X}(t)+\boldsymbol{U}(t))|\leqslant\epsilon|\boldsymbol{U}_0|.$$

从而, 当 $\boldsymbol{U}_0 \to 0$ 时, 曲线 $t \to \boldsymbol{X}(t) + \boldsymbol{U}(t)$ 将越来越逼近 $\boldsymbol{Y}(t)$. 在很多应用中, 人们直接用从变分方程得到的解 $\boldsymbol{X}(t) + \boldsymbol{U}(t)$ 来代替原来的解 $\boldsymbol{Y}(t)$; 由线性叠加原理, $\boldsymbol{U}(t)$ 线性地依赖于 $\boldsymbol{U}_0$, 这将会带来很大的方便.

例 考虑非线性方程系统

$$x' = x + y^2$$

$$y' = -y.$$

在下一章, 我们将会更详细地讨论这个系统. 而现在, 注意这个系统有一个明显的解, 也就是位于原点的平衡解 $X(t) \equiv (0,0)$. 沿这个解的变分方程为

$$\boldsymbol{U}' = \boldsymbol{DF}_0(\boldsymbol{U}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \boldsymbol{U},$$

这是一个自治的线性系统. 我们可立刻得到该方程的解为

$$\boldsymbol{U}(t) = \begin{pmatrix} x_0 \mathrm{e}^t \\ y_0 \mathrm{e}^{-t} \end{pmatrix}.$$

上面的命题就保证了, 只要 (x_0, y_0) 充分接近原点, 非线性方程经过 (x_0, y_0) 且定义在 $[-\tau, \tau]$ 上的解就可以与 $\boldsymbol{U}(t)$ 任意接近. ■

注意, 上例中的讨论是相当一般的. 任给非线性微分方程系统 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, 如果 $\boldsymbol{X}_0$ 为它的一个平衡点, 则我们就可以考虑沿这个解的变分方程. 但此时, $\boldsymbol{DF}_{\boldsymbol{X}_0}$ 是一个常值矩阵 $\boldsymbol{A}$, 从而变分方程 $\boldsymbol{U}' = \boldsymbol{AU}$ 为一个自治的线性系统. 这个系统称为在 $\boldsymbol{X}_0$ 处的**线性化系统**. 我们知道线性化系统的流为 $\exp(t\boldsymbol{A})\boldsymbol{U}_0$, 于是, 上面的命题就表明, 在平衡点附近, 一个非线性系统的相图与它的线性化系统的相图是相似的. 在下一章, 我们要将 "相似" 这一说法更准确化.

利用上一命题, 假设知道了解 $\boldsymbol{X}(t)$, 现在就可以如下来计算 $\partial\phi/\partial\boldsymbol{X}$:

定理 设 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 为一个微分方程系统, 其中 $\boldsymbol{F}$ 为 C^1 的. 令 $\boldsymbol{X}(t)$ 为该系统满足初始条件 $\boldsymbol{X}(0) = \boldsymbol{X}_0$ 的解, 它在 $t \in [\alpha, \beta]$ 上有定义; 再令 $\boldsymbol{U}(t, \boldsymbol{U}_0)$ 为沿 $\boldsymbol{X}(t)$ 的变分方程满足 $\boldsymbol{U}(0, \boldsymbol{U}_0) = \boldsymbol{U}_0$ 的解. 则

$$\boldsymbol{D}\phi_t(\boldsymbol{X}_0)\boldsymbol{U}_0 = \boldsymbol{U}(t, \boldsymbol{U}_0),$$

即 $\partial\phi/\partial\boldsymbol{X}$ 在 $\boldsymbol{U}_0$ 上的作用可由从 $\boldsymbol{U}_0$ 出发求解相应的变分方程给出.

证明 根据上一命题, 对所有的 $t \in [\alpha, \beta]$, 我们有

$$\boldsymbol{D}\phi_t(\boldsymbol{X}_0)\boldsymbol{U}_0 = \lim_{h\to 0} \frac{\phi_t(\boldsymbol{X}_0 + h\boldsymbol{U}_0) - \phi_t(\boldsymbol{X}_0)}{h} = \lim_{h\to 0} \frac{\boldsymbol{U}(t, h\boldsymbol{U}_0)}{h} = \boldsymbol{U}(t, \boldsymbol{U}_0). \quad ■$$

例 作为这些想法的一个说明, 我们来考虑微分方程 $x' = x^2$. 简单的积分表明, 满足初始条件 $x(0) = x_0$ 的解为

$$x(t) = \frac{-x_0}{x_0 t - 1}.$$

于是, 我们有

$$\frac{\partial \phi}{\partial x}(t, x_0) = \frac{1}{(x_0 t - 1)^2}.$$

另一方面, 沿 $x(t)$ 的变分方程为

$$u' = 2x(t)u = \frac{-2x_0}{x_0 t - 1} u.$$

该方程满足初始条件 $u(0) = u_0$ 的解为

$$u(t) = u_0 \left(\frac{1}{x_0 t - 1} \right)^2,$$

这正与我们期望的一样. ■

7.5 探索：数值方法

在这个探索中, 我们先描述对一阶微分方程的解进行逼近的三种不同方法. 你的任务就是对每一种方法的可行性进行评价.

这其中每个方法都包含一个迭代过程, 对一阶微分方程 $x' = f(t, x)$ 的一个解的图像上任一个选定的点 $(t_k, x(t_k))$, 我们都要找到一列点 (t_k, x_k) 去逼近它. 在每种情形, 我们都从初值 $x(0) = x_0$ 开始, 于是 $t_0 = 0$, 而 x_0 就是给定的初始值. 我们需要生成 t_k, x_k.

在三种方法中的每一种, 我们都是递归地生成 t_k：选取一个步长 Δt, 然后在 t_k 上简单地增加 Δt 就得到

$$t_{k+1} = t_k + \Delta t.$$

选取较小的 Δt(希望) 将会提高方法的精度.

从而, 为了描述每一种方法, 我们只须确定那些 x_k 的值. 在每种情形, x_{k+1} 都是 tx 平面上某一条过 (t_k, x_k) 的直线上, 位于 t_{k+1} 上方的点的 x 坐标. 于是, 只要我们给出了这条直线的斜率, x_{k+1} 也就确定了.

(1) **欧拉方法.** 此时, 沿斜率场在 (t_k, x_k) 所生成的直线移动 Δt 个时间单位就得到了 x_{k+1}. 由于 (t_k, x_k) 处的斜率就是 $f(t_k, x_k)$, 于是再增加一小步, 我们就有

$$x_{k+1} = x_k + f(t_k, x_k)\Delta t.$$

(2) **改进的欧拉方法.** 在这种方法中, 我们用两个斜率的平均来从 (t_k, x_k) 移动到 (t_{k+1}, x_{k+1}). 第一个斜率就是斜率场在 (t_k, x_k) 处的取值, 即

$$m_k = f(t_k, x_k).$$

第二个斜率则是斜率场在 (t_{k+1}, y_k) 处的取值, 其中 y_k 是在 (t_k, x_k) 处用欧拉方法得到终点, 即

$$n_k = f(t_{k+1}, y_k), \quad \text{其中 } y_k = x_k + f(t_k, x_k)\Delta t.$$

于是, 我们有

$$x_{k+1} = x_k + \left(\frac{m_k + n_k}{2}\right)\Delta t.$$

(3) **(四阶) 龙格-库塔方法.** 这是求解微分方程的最常用的方法. 对于一些特殊情形, 还有专门设计的更复杂的计算方法, 但这种方法作为一种通用算法已经用了几十年. 在这种方法中, 我们将确定四个斜率, m_k, n_k, p_k, q_k. 从 (t_k, x_k) 到 (t_{k+1}, x_{k+1}) 的这步是沿以这 4 个值取加权平均得到的斜率的直线移动得到的:

$$x_{k+1} = x_k + \left(\frac{m_k + 2n_k + 2p_k + q_k}{6}\right)\Delta t.$$

这 4 个斜率是通过以下的方式得到的:

(a) m_k 同欧拉方法中一样给出:

$$m_k = f(t_k, x_k).$$

(b) 沿在 (t_k, x_k) 处的斜率场线移动半个步长得到中间点 $(t_k + (\Delta t)/2, y_k)$, n_k 就是斜率场在该点的斜率, 即

$$n_k = f\left(t_k + \frac{\Delta t}{2}, y_k\right), \quad \text{其中 } y_k = x_k + m_k\frac{\Delta t}{2}.$$

(c) 沿过 (t_k, x_k) 点的斜率为 n_k (而不是 m_k) 的直线移动半个步长得到另一个中间点 $(t_k + (\Delta t)/2, z_k)$, p_k 就是斜率场在该点的斜率, 即

$$p_k = f\left(t_k + \frac{\Delta t}{2}, z_k\right), \quad \text{其中 } z_k = x_k + n_k\frac{\Delta t}{2}.$$

(d) 用过 (t_k, x_k) 点的斜率为 p_k 的直线得到点 (t_{k+1}, w_k), q_k 就是斜率场在该点的斜率, 即

$$q_k = f(t_{k+1}, w_k), \quad \text{其中 } w_k = x_k + p_k\Delta t.$$

在这个探索中, 你们的目的就是, 通过估计在几个例子中实现这些程序所产生的误差, 从而比较 3 种方法的效果. 我们建议你们用一个电子表格程序来做这些冗长的计算.

(1) 首先, 为了确信你们理解上面 3 种方法相当精炼的描述, 在其中的每种情形, 请在 tx 平面上画图示意从 (t_k, x_k) 移动到 (t_{k+1}, x_{k+1}) 的过程.

(2) 然后, 我们来研究各种方法对一个特别简单的微分方程 $x' = x$ 是如何工作的.

(a) 首先, 找出这个方程满足初始条件 $x(0) = 1$ 的显式解 $x(t)$(此时, 这是来自于数学系的免费礼物).

(b) 然后, 用欧拉方法来逼近 $x(1) = \mathrm{e}$ 这个值, 步长为 $\Delta t = 0.1$. 即利用 $\Delta t = 0.1$, 从 $t_0 = 0, x_0 = 1$ 出发, 对于 $k = 1, \cdots, 10$, 递归地确定出 t_k, x_k.

(c) 将步长减半, 即取 $\Delta t = 0.05$, 重复上一步.

(d) 还是用欧拉方法来逼近 $x(1)$, 不过这次将步长除以 5, 也就是取 $\Delta t = 0.01$.

(e) 现在, 使用改进的欧拉方法以及同样的那些步长重复上面的 3 步.

(f) 然后, 使用龙格-库塔方法再重做.

(g) 此时, 你就得到了 $x(1) = \mathrm{e}$ 的 9 个不同的逼近值, 每种方法 3 个. 对每种情形计算误差. 在计算误差时, 取 $\mathrm{e} = 2.718\,281\,828\,452\,353\,602\,87\cdots$ 作为其精确值.

(h) 最后计算, 当步长从 0.1 变到 0.05 以及从 0.05 变到 0.01 时, 误差是如何变化的. 即如果记 ρ_Δ 为步长为 Δ 时所产生的误差, 计算 $\rho_{0.1}/\rho_{0.05}$ 和 $\rho_{0.05}/\rho_{0.01}$.

(3) 现在重复前面的探索, 这次是对非自治方程 $x' = 2t(1+x^2)$. 利用值 $\tan 1 = 1.557\,407\,724\,654\cdots$.

(4) 当把步长缩小到 1/2 或 1/5 后, 讨论误差是如何变化的. 特别地, 为什么龙格-库塔方法称为“四阶”方法?

习　题

1. 对于下面的每个初值问题, 写出毕卡迭代法的开始几项. 如果可能, 找出显式解, 并描述解的定义域.
   (a) $x' = x + 2$; $x(0) = 2$.
   (b) $x' = x^{4/3}$; $x(0) = 0$.
   (c) $x' = x^{4/3}$; $x(0) = 1$.
   (d) $x' = \cos x$; $x(0) = 0$.
   (e) $x' = 1/(2x)$; $x(1) = 1$.
2. 令 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵. 证明用毕卡迭代法求解 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}, \boldsymbol{X}(0) = \boldsymbol{X}_0$ 可以给出解 $\exp(t\boldsymbol{A})\boldsymbol{X}_0$.

3. 将毕卡迭代法运用到二阶初值问题

$$x'' = -4x; \quad x(0) = 0, \quad x'(0) = 2$$

对应的一阶系统, 从而得到 $\sin 2t$ 的泰勒级数.

4. 对线性非自治微分方程系统验证线性叠加原理.
5. 考虑一阶方程 $x' = x/t$. 对于满足 $x(0) = 0$ 的解你能说些什么? 对于 $x(0) = a \neq 0$ 呢?
6. 讨论初值问题 $x' = x^a, x(0) = 0$ 的解的存在唯一性, 其中 $a > 0$.
7. 设 $\boldsymbol{A}(t)$ 为一族连续的 $n \times n$ 矩阵, 且 $\boldsymbol{P}(t)$ 为初值问题 $\boldsymbol{P}' = \boldsymbol{A}(t)\boldsymbol{P}, \boldsymbol{P}(0) = \boldsymbol{P}_0$ 的矩阵解. 证明

$$\det \boldsymbol{P}(t) = (\det \boldsymbol{P}_0) \exp\left(\int_0^t \mathrm{Tr}\boldsymbol{A}(s)\mathrm{d}s\right).$$

8. 构造 $\mathbb{R}$ 上的一个一阶微分方程的例子, 使得它对任何初值问题都没有解.
9. 构造一个依赖于参数 a 的微分方程的例子, 使得它的某些解不连续依赖于 a.

第 8 章　非线性系统的平衡点

为了避免出现上一章所遇到的一些技术性困难, 从现在起, 除非特别申明, 总假设我们的微分方程为 C^∞ 的. 这意味着, 对所有的 k, 微分方程的右端都是 k 次连续可微的. 这至少可以保证我们定理中的假设条件是最少的.

我们已经看到, 写出非线性微分方程系统的显式解往往是不可能的. 当我们有平衡解时则是一个例外. 只要我们可以求解相应的代数方程, 我们就可以将平衡点明确地写出来. 常常, 对于一些特定的非线性系统, 这些就是它们最重要的解. 更重要的是, 在已经有了有关线性系统的大量的工作后, 我们通常可以利用**线性化**的技巧, 来确定平衡点附近的解行为. 在本章我们将详细地描述这一过程.

8.1　一些用作说明的例子

在本节, 我们考虑几个微分方程的平面非线性系统, 原点都是它们的平衡点. 我们的目的是看出非线性系统在原点附近的解与它们的**线性化**系统的解是相似的, 至少在某些情形是如此.

作为第一个例子, 考虑系统

$$\begin{aligned} x' &= x + y^2 \\ y' &= -y. \end{aligned}$$

它仅有一个位于原点的平衡点. 为了画出附近的解, 我们注意到, 如果 y 很小, 则 y^2 就更小. 于是至少在原点附近, 微分方程 $x' = x + y^2$ 与 $x' = x$ 非常接近. 在 7.4 节中, 我们已经证明了这个系统的流在原点附近与它的线性化系统 $\boldsymbol{X}' = \boldsymbol{DF}_0\boldsymbol{X}$ 的流很“接近”. 由此想到, 我们可以不考虑线性化系统, 而是考虑简单地扔掉高阶项后的系统

$$\begin{aligned} x' &= x \\ y' &= -y. \end{aligned}$$

显然, 我们可以立刻解出这个系统. 于是我们得到在原点的一个鞍点, 它的稳定线为 y 轴, 而不稳定线为 x 轴.

现在让我们回到原来的非线性系统. 很幸运, 我们也可以将这个系统显式地解出来. 从第二个方程 $y' = -y$ 可得 $y(t) = y_0\mathrm{e}^{-t}$, 将它代入第一个方程, 可得

$$x' = x + y_0^2\mathrm{e}^{-2t}.$$

这是一个一阶非自治方程, 像微积分中一样, 我们 “猜” 它有一个形如 ce^{-2t} 的特解, 这样就可以确定它的解. 事实上, 将猜测的这个形式代入方程就可以得到一个特解:

$$x(t) = -\frac{1}{3}y_0^2\mathrm{e}^{-2t}.$$

于是, 很容易验证, 任何如下形式的函数

$$x(t) = c\mathrm{e}^t - \frac{1}{3}y_0^2\mathrm{e}^{-2t}$$

都是方程的一个解. 从而原系统的通解为

$$\begin{aligned} x(t) &= \left(x_0 + \frac{1}{3}y_0^2\right)\mathrm{e}^t - \frac{1}{3}y_0^2\mathrm{e}^{-2t} \\ y(t) &= y_0\mathrm{e}^{-t}. \end{aligned}$$

如果 $y_0 = 0$, 我们可找到一个直线解 $x(t) = x_0\mathrm{e}^t, y(t) = 0$, 这与线性情形的解是相同的. 然而, 与线性情形不同的是, y 轴不再是趋于原点的一个解. 事实上, 沿 y 轴的向量场为 $(y^2, -y)$, 它与 y 轴不相切, 其实, 所有的非零向量都是指向该坐标轴的右侧.

另一方面, 有一条过原点的曲线, 上面的解都趋向 $(0,0)$. 考虑 $\mathbb{R}^2$ 中的曲线 $x + \frac{1}{3}y^2 = 0$. 假设 (x_0, y_0) 在这条曲线上, 令 $(x(t), y(t))$ 是满足这个初始条件的解. 由于 $x_0 + \frac{1}{3}y_0^2 = 0$, 这个解就是

$$\begin{aligned} x(t) &= -\frac{1}{3}y_0^2\mathrm{e}^{-2t} \\ y(t) &= y_0\mathrm{e}^{-t}. \end{aligned}$$

注意, 对所有的 t, 我们都有 $x(t) + \frac{1}{3}(y(t))^2 = 0$, 从而对所有的时间, 这个解都停留在曲线上. 而且, 当 $t \to \infty$ 时, 这个解趋于平衡点. 这样, 我们就找到了过原点的一条**稳定曲线**,它上面的所有解都趋于 $(0,0)$. 注意, 这条曲线在原点与 y 轴相切 (见图 8.1).

我们能否真的 “扔掉” 系统的非线性项呢? 下面我们将看到, 这个问题的回答是: 得看情况. 然而, 在现在的情形, 这样做是完全合法的, 这是因为我们能找到一个变量替换将原来的系统真的转化成它的线性系统.

为了说明这一点, 我们引入新变量 u, v 如下

$$\begin{aligned} u &= x + \frac{1}{3}y^2 \\ v &= y. \end{aligned}$$

则在新坐标下, 系统变成

$$u' = x' + \frac{2}{3}yy' = x + \frac{1}{3}y^2 = u$$

$$v' = y' = -y = -v.$$

即, 非线性的变量替换 $F(x,y) = \left(x + \frac{1}{3}y^2, y\right)$ 将原来的非线性系统转化成了一个线性系统, 事实上, 就是上面的线性化系统.

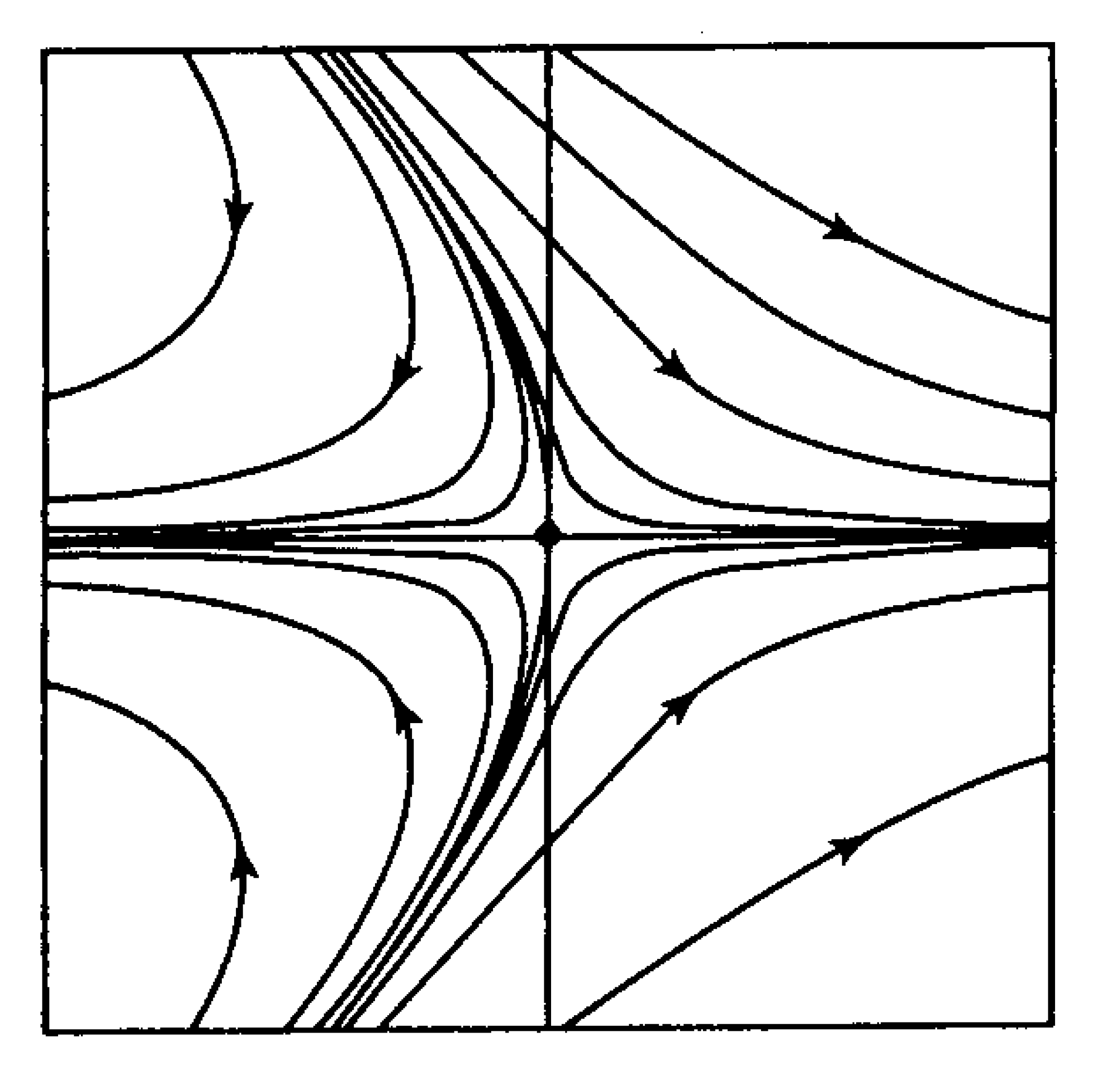


图 8.1 $x' = x + y^2, y' = -y$ 的相平面. 注意它的稳定曲线与 y 轴相切

例 一般地, 像上例那样将一个非线性系统转换成一个线性系统是不可能的, 这是因为非线性项几乎总是会对系统在远离原点 (系统的平衡点) 的地方产生巨大的改变. 例如, 考虑非线性系统

$$x' = \frac{1}{2}x - y - \frac{1}{2}(x^3 + y^2x)$$

$$y' = x + \frac{1}{2}y - \frac{1}{2}(y^3 + x^2y).$$

我们在原点处仍然有一个平衡点. 现在的线性化系统为

$$\boldsymbol{X}' = \begin{pmatrix} \frac{1}{2} & -1 \\ 1 & \frac{1}{2} \end{pmatrix} \boldsymbol{X},$$

其特征值为 $\frac{1}{2} \pm \mathrm{i}$. 容易验证, 这个线性系统的所有解都按反时针方向盘旋地远离原点, 且趋于 ∞.

求解上面的非线性系统看起来很困难. 然而, 如果换成极坐标, 方程会变得非

常简单. 通过计算可得

$$r'\cos\theta - r(\sin\theta)\theta' = x' = \frac{1}{2}(r-r^3)\cos\theta - r\sin\theta$$
$$r'\sin\theta + r(\cos\theta)\theta' = y' = \frac{1}{2}(r-r^3)\sin\theta + r\cos\theta,$$

在上面的式子中比较 $\cos\theta$ 和 $\sin\theta$ 的系数可得

$$r' = r(1-r^2)/2$$
$$\theta' = 1.$$

由于去掉了方程的耦合, 现在我们可将系统显式地解出. 为了做得更好些, 我们将以一种更几何化的方式进行. 从方程 $\theta'=1$ 可得, 所有的非零解都绕着原点反时针方向盘旋. 从第一个方程可以看出, 解并不会盘旋地趋于 ∞. 事实上, 当 $r=1$ 时, 我们有 $r'=0$, 于是所有从单位圆周上出发的解永远逗留在那儿, 而且沿圆周作周期运动. 由于当 $0<r<1$ 时, $r'>0$, 我们可以得出, 单位圆周内部的非零解都盘旋地离开原点, 并且趋向单位圆周. 由于当 $r>1$ 时, $r'<0$, 单位圆周外部的解都盘旋地趋向单位圆周 (见图 8.2). ■

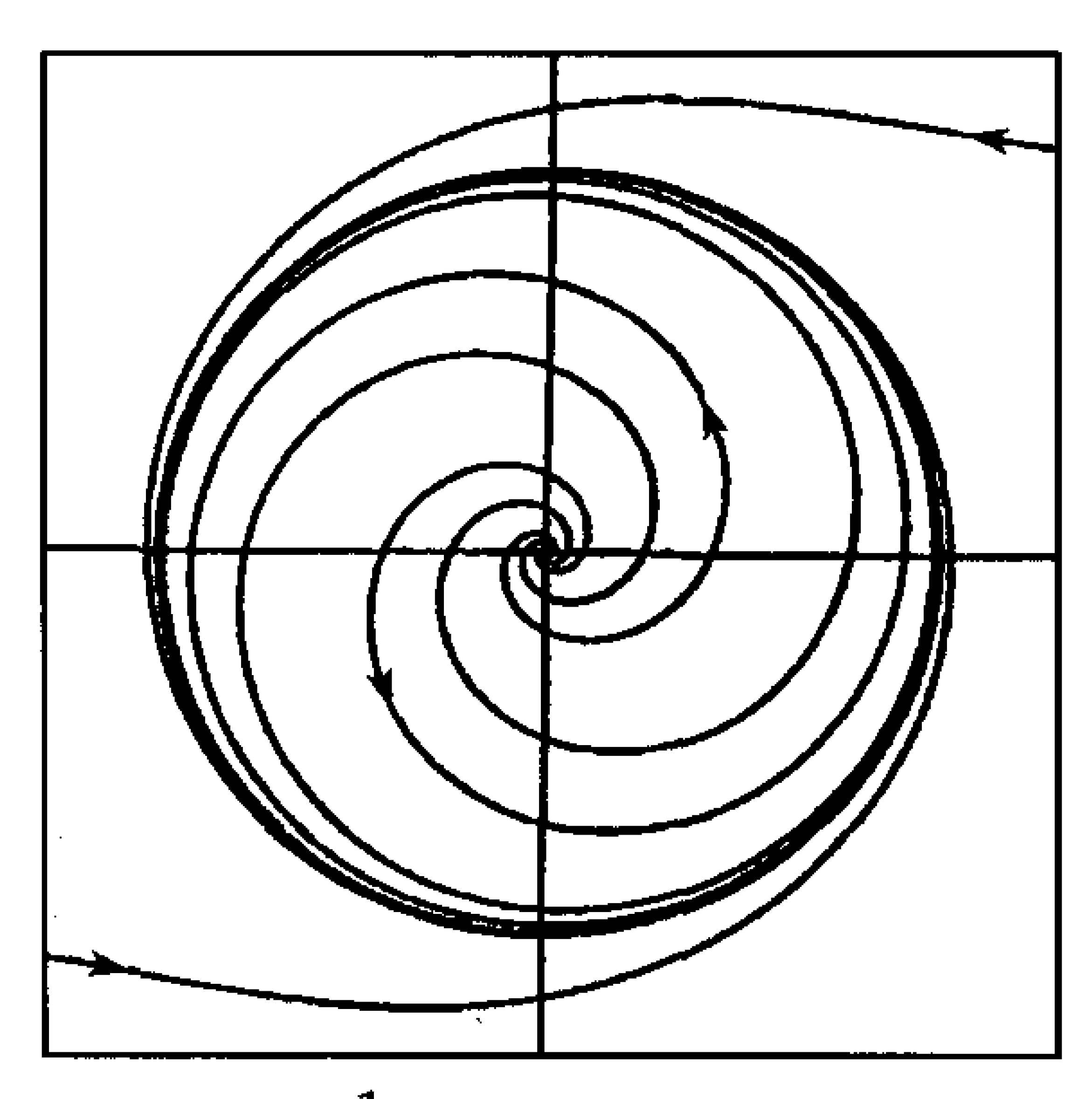


图 8.2　$r'=\frac{1}{2}(r-r^3), \theta'=1$ 的相平面

在这个例子中, 没办法找到一个整体的坐标变换将系统化成线性形式, 其理由是, 线性系统都不会像这样盘旋地趋向圆周. 然而在原点附近仍然有可能.

为了看出这点, 首先注意到, 当 $0<r_0<1$ 时, 在半径为 r_0 的圆周上, 非线性向量场指向这个圆周的外部. 这是由于在任何一个这样的圆周上, 都有 $r'=r_0(1-r_0^2)/2>0$. 从而, 当时间向后时, 非线性系统的所有解都趋于原点, 事实上是盘旋地趋于原点.

利用这个事实, 我们可以在圆盘 $r\leqslant r_0$ 内定义线性和非线性系统之间的一个共轭, 其方式类似于第 4 章. 记 ϕ_t 为非线性系统的流. 在极坐标中, 上面的线性化

系统成为

$$r' = r/2$$
$$\theta' = 1.$$

记 ψ_t 为这个系统的流. 在时间向后时, 这个线性系统的所有解也都趋于原点. 现在在由 $r<1$ 定义的圆盘 D 内定义这两个流之间的一个共轭. 固定 $r_0, 0<r_0<1$. 对于 D 内的任一点 $(r,\theta), r>0$, 存在唯一的 $t=t(r,\theta)$ 使得 $\phi_t(r,\theta)$ 属于圆周 $r=r_0$. 然后, 我们定义函数

$$h(r,\theta) = \psi_{-t} \circ \phi_t(r,\theta),$$

其中, $t=t(r,\theta)$. 我们还规定 h 将原点带到原点. 则可以直接验证, 对于 D 内任一点 (r,θ), 都成立

$$h \circ \phi_s(r,\theta) = \psi_s \circ h(r,\theta),$$

从而, h 就给出了非线性系统和线性系统之间的一个共轭. 同时, 容易验证, h 将 D 映射到整个 $\mathbb{R}^2$.

于是, 我们发现, 虽然并不是总能将一个系统**整体**线性化, 但时常能够实现**局部**线性化. 很遗憾, 这样的局部线性化可能对解的行为提供不了任何有用的信息.

例 现在, 考虑系统

$$x' = -y + \epsilon x(x^2+y^2)$$
$$y' = x + \epsilon y(x^2+y^2).$$

这里 ϵ 为一个参数, 既可取正也可取负.

它的线性化系统为

$$x' = -y$$
$$y' = x,$$

由此可见, 原点为一个中心, 而且所有的解都在以原点为中心的圆周上, 以单位角速度反时针方向旋转.

非线性系统却几乎完全不一样. 在极坐标下, 非线性系统成为

$$r' = \epsilon r^3$$
$$\theta' = 1.$$

于是, 当 $\epsilon>0$ 时, 所有的解都盘旋地离开原点, 而当 $\epsilon<0$ 时, 所有的解都盘旋地趋于原点. 增加的非线性项, 不管它在原点附近是多么的小, 都明显地改变了线性化系统的相图; 从而, 我们不能用线性化方法确定系统在平衡点附近的行为. ■

例 现在考虑最后一个例子:

$$x' = x^2$$
$$y' = -y.$$

该系统唯一的平衡解为原点. 所有其他的解 (除了 y 轴上的解) 都向右运动, 并且趋向 x 轴. 在 y 轴上, 解都沿该直线趋于原点. 于是, 其相图如图 8.3 所示. ■

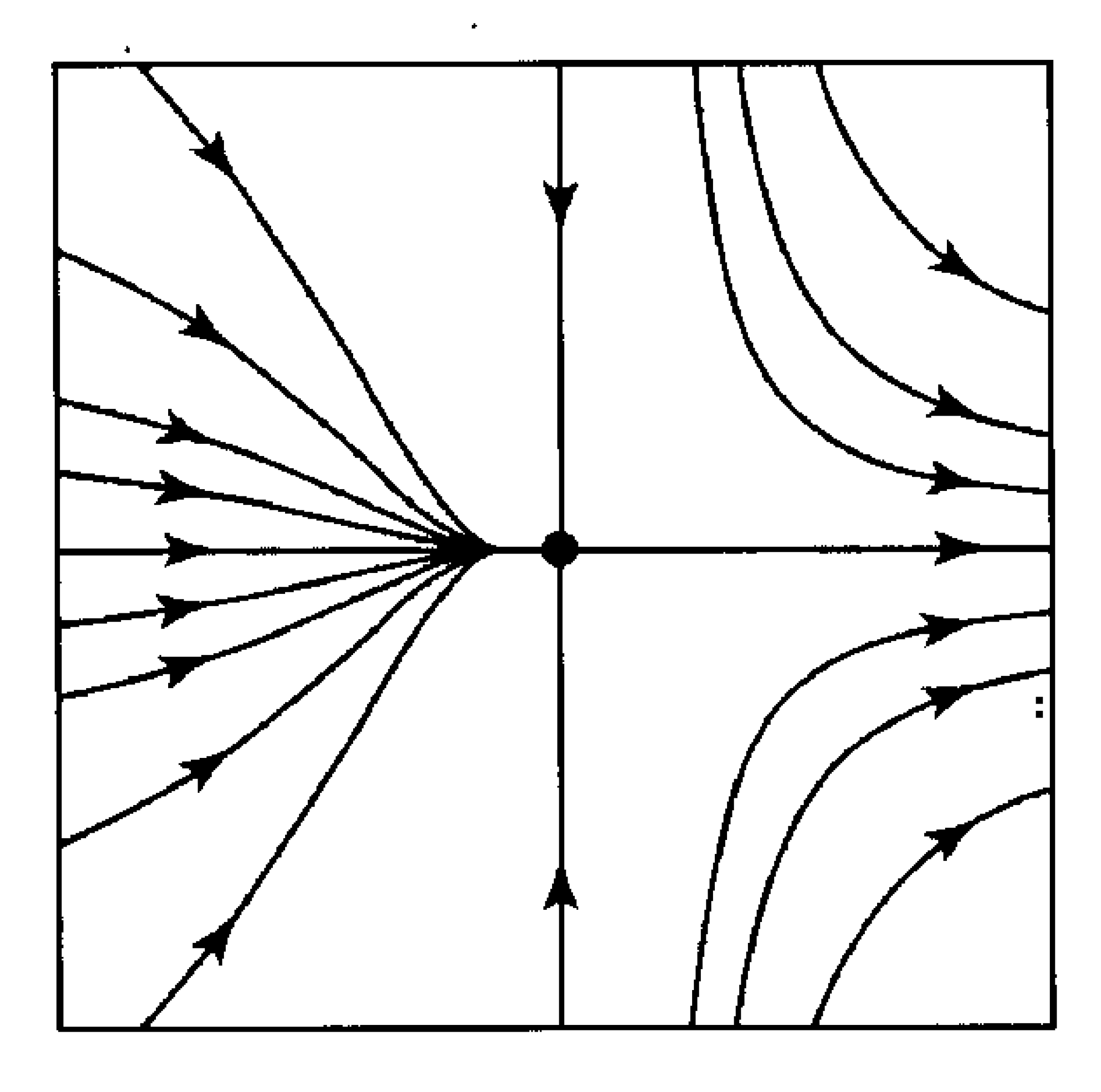


图 8.3　$x' = x^2, y' = -y$ 的相平面

该系统的线性化系统为

$$x' = 0$$
$$y' = -y,$$

x 轴上的所有的点都是这个系统的平衡点, 而所有其他的解都位于 $x = c$(c 为常数) 的竖线上. 可见, 它与图 8.3 的图形十分的不同.

这里, 以及上例中的问题在于, 线性化系统在原点的平衡点不是双曲的. 当一个平面线性系统有一个零特征值或中心时, 增加的非线性项常常会将相图完全改变.

8.2　非线性的汇点和源点

在上节的例子中我们看到, 只有当线性化系统为双曲时 (也就是没有一个特征值的实部为零), 平面非线性系统在平衡点附近的解行为才会与其线性化系统相似. 在本节, 通过考虑汇点这一特殊情形, 我们开始描述一般的非线性系统中的双曲平衡点. 为了简单起见, 我们将只对平面情形给出结果的证明, 当然, 所有的这些结果在 $\mathbb{R}^n$ 中也都成立.

令 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, 并假设 $\boldsymbol{F}(\boldsymbol{X}_0) = 0$. 记 $\boldsymbol{DF}_{\boldsymbol{X}_0}$ 为 $\boldsymbol{F}$ 在 $\boldsymbol{X}_0$ 处的雅可比矩阵. 则, 像第 7 章一样, 线性微分方程系统

$$\boldsymbol{Y}' = \boldsymbol{DF}_{\boldsymbol{X}_0}\boldsymbol{Y}$$

称为**$\boldsymbol{X}_0$ 处的线性化系统**. 注意, 如果 $\boldsymbol{X}_0 = \boldsymbol{0}$, 线性化系统可以通过简单地扔掉 $\boldsymbol{F}$ 中的非线性项得到, 我们在上节就是这么做的.

为了与我们在线性系统的工作进行类比, 我们称一个非线性系统的一个平衡点 $\boldsymbol{X}_0$ 为**双曲的**, 如果 $\boldsymbol{DF}_{\boldsymbol{X}_0}$ 的所有特征值都有非零实部.

下面我们专门讨论平面系统的平衡点, 而且假设相应的线性化系统在原点处为一汇点. 假设我们的系统为

$$x' = f(x,y)$$
$$y' = g(x,y),$$

其中 $f(x_0,y_0) = 0 = g(x_0,y_0)$. 如果我们做坐标变换 $u = x - x_0, v = y - y_0$, 则新系统以 $(0,0)$ 为一个平衡点. 因而, 我们不妨在开始就假设 $x_0 = y_0 = 0$. 然后再做一个线性的坐标变换将线性化系统化为标准形. 为了简单起见, 我们在开始就假设线性化系统有不同的特征值 $-\lambda < -\mu < 0$. 于是, 在经过这些坐标变换之后, 我们的系统就变成了

$$x' = -\lambda x + h_1(x,y)$$
$$y' = -\mu y + h_2(x,y),$$

其中 $h_j = h_j(x,y)$ 包含所有的 "高阶项", 即, h_j 含有它的泰勒展式中的 x, y 的二阶或更高阶项. 等价地, 我们有

$$\lim_{(x,y)\to(0,0)} \frac{h_j(x,y)}{r} = 0,$$

其中 $r^2 = x^2 + y^2$.

此时, 线性化系统就是

$$x' = -\lambda x$$
$$y' = -\mu y.$$

对于这个线性化系统, 我们知道向量场总是指向中心在原点, 半径为 r 的圆周的内部. 事实上, 如果求这个线性向量场与圆周的外法向量 (x, y) 的点乘, 就有, 对任何的非零向量 (x,y),

$$(-\lambda x, -\mu y) \cdot (x,y) = -\lambda x^2 - \mu y^2 < 0.$$

正如我们在第 4 章中所见, 这将使得所有解的径向分量严格递减, 从而所有解都趋于原点.

对非线性系统也会发生同样的事情, 至少在 $(0,0)$ 附近会如此. 记 $q(x,y)$ 为向量场与 (x,y) 的点乘. 我们有

$$\begin{aligned} q(x,y) &= (-\lambda x + h_1(x,y), -\mu y + h_2(x,y)) \cdot (x,y) \\ &= -\lambda x^2 + x h_1(x,y) - \mu y^2 + y h_2(x,y) \\ &= -\mu(x^2+y^2) + (\mu - \lambda)x^2 + x h_1(x,y) + y h_2(x,y) \\ &\leqslant -\mu r^2 + x h_1(x,y) + y h_2(x,y), \end{aligned}$$

最后一个不等式是因为 $(\mu-\lambda)x^2\leqslant 0$. 于是, 我们就有

$$\frac{q(x,y)}{r^2}\leqslant -\mu+\frac{xh_1(x,y)+yh_2(x,y)}{r^2}.$$

当 $r\to 0$ 时, 上式的右边趋于 $-\mu$. 由此可得, 至少在充分靠近原点时, $q(x,y)$ 为负的. 从而, 非线性向量场指向中心在原点, 半径很小的圆周的内部, 因而当初始条件取在这些圆周的内部时, 相应的解必定趋于原点. 于是, 这就解释了, 为什么我们像线性情形一样称这种类型的平衡点为**汇点**.

可以直接验证, 当线性化系统的特征值为 $\alpha\pm i\beta, \alpha<0, \beta\neq 0$ 时, 同样的结果也成立. 在重负特征值情形, 我们首先需要作坐标变换使得线性化系统成为

$$\begin{aligned}x'&=-\lambda x+\epsilon y\\ y'&=-\lambda y,\end{aligned}$$

其中, ϵ 充分小. 在第 4 章中我们已经说明了如何处理这种情况. 于是, 同样地有, 该向量场指向半径充分小的圆周的内部.

现在我们就能够将一个非线性系统的流在一个双曲平衡点 (汇点) 附近共轭到它的线性化系统的流. 事实上, 上节第二个例子的讨论可以几乎不作变动就用到现在的情形. 类似地, 在一个双曲源点附近, 非线性系统也是共轭于相应的线性化系统.

这个结果是下面更一般的定理的一个特殊情形.

线性化定理 假设 n 维系统 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 在 $\boldsymbol{X}_0$ 处有一个双曲平衡点. 则, 在 $\boldsymbol{X}_0$ 的一个邻域内, 非线性系统的流共轭于其线性化系统的流. ■

由于当特征值既有正实部又有负实部时, 证明需要用到超出本书范围的分析技巧, 因而我们不打算在这里证明这个定理.

8.3 鞍　　点

现在我们来看线性化系统在 $\mathbb{R}^2$ 的原点处为一鞍点的平衡点情形. 像上节一样, 我们不妨假设系统有如下的形式:

$$\begin{aligned}x'&=\lambda x+f_1(x,y)\\ y'&=-\mu y+f_2(x,y),\end{aligned}$$

其中 $-\mu<0<\lambda$, 且当 $r\to 0$ 时, $f_j(x,y)/r$ 趋于 0. 像线性系统情形一样, 我们称这种类型的平衡点为一个**鞍点**.

对于线性化系统, y 轴为稳定线, 当 $t\to\infty$ 时, 此线上的所有解都趋于原点. 类似地, x 轴为不稳定线. 正如在 8.1 节中所看到, 我们不能指望这些稳定和不稳

定直线在非线性情形还保持. 尽管如此, 确实还存在一对过原点的曲线具有类似的性质.

记 $W^s(0)$ 为一些初始条件构成的集合, 从这些初始条件出发的解当 $t \to \infty$ 时趋于原点. 记 $W^u(0)$ 为另一些初始条件构成的集合, 从这些初始条件出发的解当 $t \to -\infty$ 时趋于原点. $W^s(0)$ 和 $W^u(0)$ 分别称为**稳定曲线**和**不稳定曲线**.

下面的定理表明非线性鞍点附近解的行为与线性情形非常相似.

稳定曲线定理 假设系统

$$\begin{aligned} x' &= \lambda x + f_1(x, y) \\ y' &= -\mu y + f_2(x, y) \end{aligned}$$

满足 $-\mu < 0 < \lambda$, 且当 $r \to 0$ 时, $f_j(x, y)/r$ 趋于 0. 则存在 $\epsilon > 0$ 以及定义在 $|y| < \epsilon$ 上的一条曲线 $x = h^s(y)$ 满足 $h^s(0) = 0$. 进一步,

(1) 初始条件在这条曲线上的所有的解对所有的 $t \geqslant 0$ 仍然在这条曲线上, 而且在 $t \to \infty$ 时趋于原点;

(2) 曲线 $x = h^s(y)$ 经过原点时切于 y 轴;

(3) 初始条件位于以原点为中心、ϵ 为半径的圆盘内的所有其他的解在时间增加时都将离开该圆盘. ■

我们依次给出几点注记. 曲线 $x = h^s(y)$ 称为 0 处的**局部稳定曲线**. 当位于局部稳定曲线上的解在时间向后时就得到完整的稳定曲线 $W^s(0)$. 事实上, 函数 $h^s(y)$ 在所有点处都是 C^∞, 但我们并不打算在此证明这一点.

类似的不稳定曲线定理将为我们提供形如 $y = h^u(x)$ 的**局部不稳定曲线**. 这条曲线在原点与 x 轴相切. 曲线上的所有解当 $t \to -\infty$ 时都趋于原点.

我们先给出稳定曲线定理证明的一个简单的梗概. 考虑由直线 $|x| = \epsilon, |y| = \epsilon$ 围成的正方形, 其中 $\epsilon > 0$ 充分小. 线性系统 $x' = \lambda x, y' = -\mu y$ 的向量场在正方形的上下两个边界 $y = \pm\epsilon$ 的内部指向正方形的内部, 由于非线性系统和这个线性系统很接近, 因而非线性系统也具有这个性质. 类似地, 非线性向量场在左右两个边界上指向正方形的外部.

现在考虑位于上边界 $y = \epsilon$ 上的初始条件, 其中的有些解将从正方形的左边离开, 而另外一些则从右边离开, 但解不可能既从左边又从右边离开, 从而这些解的集合是不相交的. 于是, 一定有一些初始条件使得从它们出发的解根本就不会离开. 我们将首先证明这些不离开的解将趋于原点. 其次, 我们将证明在上下两个边界上都分别只有一个初始条件使得它的解的行为是这样的. 最后, 我们将证明这个解在一个形如 $x = h^s(y)$ 的函数的图像上, 并且函数 $x = h^s(y)$ 具有所要的性质.

下面, 我们将证明的细节补上. 记 B_ϵ 为由 $x = \pm\epsilon, y = \pm\epsilon$ 围成的正方形, 记 $S_\epsilon^\pm$ 为 B_ϵ 的上下边界. 记 C_M 为 B_ϵ 内部由 $|y| \geqslant M|x|$ 所定义的锥形区域. 这里我

们只关心 C_M 的边界的斜率 $\pm M$ 的绝对值很大的情形 (见图 8.4).

引理 任给 $M>0$, 存在 $\epsilon>0$ 使得在 $C_M\cap B_\epsilon$ 的边界点 (当然, 原点除外) 上, 非线性向量场指向 C_M 的外部.

证明 任给 $M>0$, 选取 $\epsilon>0$ 使得对所有的 $(x,y)\in B_\epsilon$, 都有

$$|f_1(x,y)|\leqslant\frac{\lambda}{2\sqrt{M^2+1}}r.$$

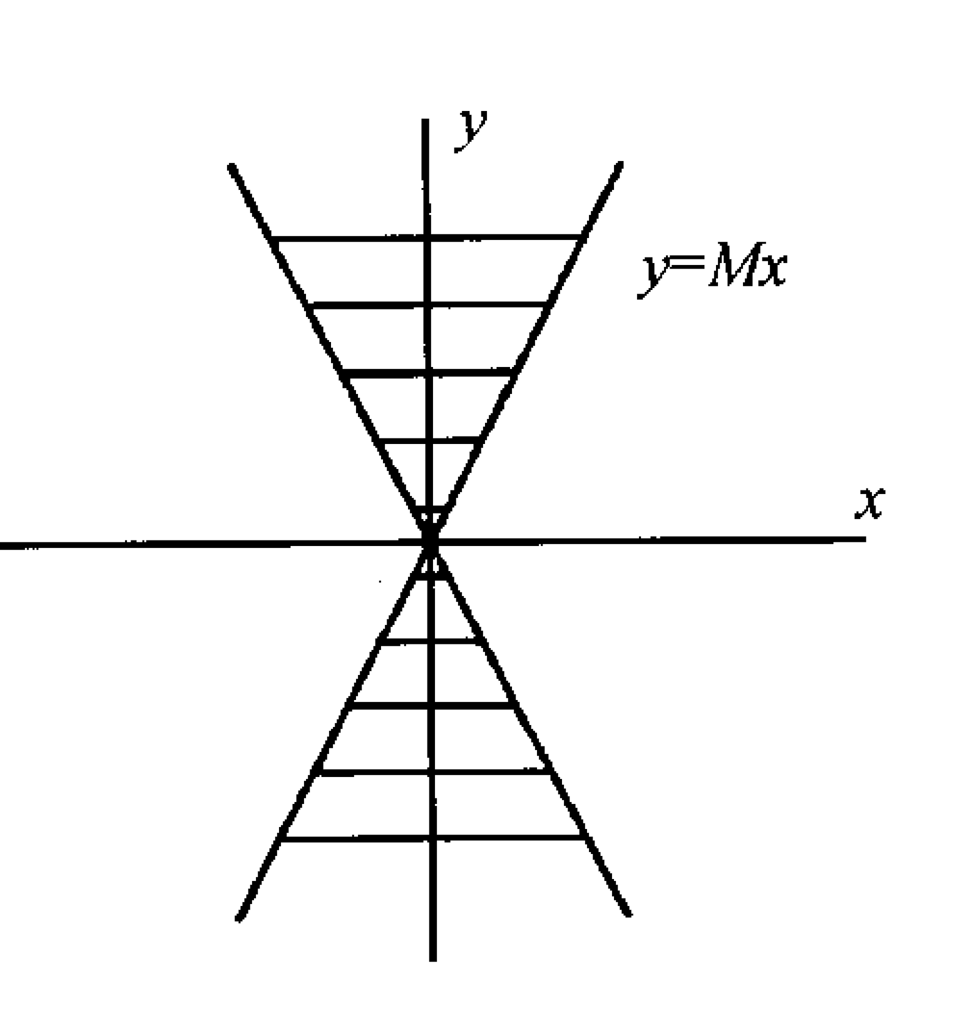


图 8.4 锥 C_M

下面假设 $x>0$. 在 C_M 的右边界上, 我们有

$$\begin{aligned}x' &= \lambda x+f_1(x,Mx)\\ &\geqslant \lambda x-|f_1(x,Mx)|\\ &\geqslant \lambda x-\frac{\lambda}{2\sqrt{M^2+1}}r\\ &= \lambda x-\frac{\lambda}{2\sqrt{M^2+1}}\left(x\sqrt{M^2+1}\right)\\ &= \frac{\lambda}{2}x>0.\end{aligned}$$

于是, 在锥形边界的这个边上有 $x'>0$.

类似地, 如果 $y>0$, 必要时取更小的 $\epsilon>0$ 使得在 $y>0$ 的 C_M 的边界上, 有 $y'<0$. 事实上, 只要选取 ϵ 满足

$$|f_2(x,y)|\leqslant\frac{\mu}{2\sqrt{M^2+1}}r$$

即可. 于是, 在 C_M 的位于第一象限的边界上, 我们已经证明向量场指向右下, 从而指向 C_M 的外部. 类似的推理可以证明在 C_M 的其他边界上, 向量场也是指向 C_M 的外部. 这就证明了该引理. ■

从这个引理可得, 存在 $S_\epsilon^\pm\cap C_M$ 中的一个初始条件集合使得其解最终从 C_M 的右边离开, 以及 $S_\epsilon^\pm\cap C_M$ 中的另一个初始条件集合使得其解最终从 C_M 的左边离开. 由于解对初始条件的连续性 (见 7.3 节), 这两个集合都是开集. 下面我们证明, 这两个集合中的每一个事实上都是单一的开区间.

记 C_M^+ 为 C_M 位于 x 轴上方的部分, 而 C_M^- 为 C_M 位于 x 轴下方的部分.

引理 假设 $M>1$, 则存在 $\epsilon>0$ 使得在 C_M^+ 中 $y'<0$, 在 C_M^- 中 $y'>0$.

证明 在 C_M^+ 中, 我们有 $|Mx|\leqslant y$, 从而

$$r^2\leqslant\frac{y^2}{M^2}+y^2,$$

也即

$$r \leqslant \frac{y}{M}\sqrt{1+M^2}.$$

像上一引理一样, 选取 ϵ 使得对所有的 $(x,y)\in B_\epsilon$ 都有

$$|f_2(x,y)| \leqslant \frac{\mu}{2\sqrt{M^2+1}}r.$$

于是, 在 C_M^+ 中, 我们就有

$$\begin{aligned} y' &\leqslant -\mu y + |f_2(x,y)| \\ &\leqslant -\mu y + \frac{\mu}{2\sqrt{M^2+1}}r \\ &\leqslant -\mu y + \frac{\mu}{2M}y \\ &\leqslant -\frac{\mu}{2}y, \end{aligned}$$

最后一个不等式成立是因为 $M>1$. 这就对 C_M^+ 证明了结果. 对 C_M^- 的证明是类似的. ■

从这个结果, 我们可以看出, 从 $S_\epsilon^+\cap C_M$ 出发的解, 如果仍然在 C_M^+ 中, 则沿 y 方向递减. 特别地, 没有解会始终待在 C_M^+ 中, 除非解趋于原点. 由存在唯一性定理, S_ϵ^+ 中从右边 (或左边) 离开的点必定形成一个单一的开区间. 因而这两个区间在 S_ϵ^+ 中的余集为一个非空闭区间, 其上的解始终不离开 C_M^+, 因而当 $t\to\infty$ 时趋于 0. 在 C_M^- 中, 我们有类似的结果.

现在, 我们断言, 在 $S_\epsilon^\pm$ 中解趋于 0 的初始条件区间实际上是一个单点. 为了看出这一点, 回想在 7.4 节中讨论的这个系统的变分方程, 这是一个非自治线性系统

$$\boldsymbol{U}' = \boldsymbol{DF}_{\boldsymbol{X}(t)}\boldsymbol{U},$$

其中 $\boldsymbol{X}(t)$ 为非线性系统的一个给定解,

$$\boldsymbol{F}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda x + f_1(x,y) \\ -\mu y + f_2(x,y) \end{pmatrix}.$$

我们在第 7 章中已经证明, 任给 $\tau>0$, 如果 $|\boldsymbol{U}_0|$ 充分小, 则当 $0\leqslant t\leqslant\tau$ 时, 变分方程满足 $\boldsymbol{U}(0)=\boldsymbol{U}_0$ 的解 $\boldsymbol{U}(t)$ 与非线性系统满足 $\boldsymbol{Y}(0)=\boldsymbol{X}_0+\boldsymbol{U}_0$ 的解 $\boldsymbol{Y}(t)$ 彼此都很接近.

现在令 $\boldsymbol{X}(t)=(x(t),y(t))$ 为该系统从不离开 C_M 的一个解. 假设 $x(0)=x_0, y(0)=\epsilon$.

引理　令 $\boldsymbol{U}(t)=(u_1(t),u_2(t))$ 为变分方程

$$\boldsymbol{U}'=\boldsymbol{DF}_{\boldsymbol{X}(t)}\boldsymbol{U}$$

满足 $|u_2(0)/u_1(0)|\leqslant 1$ 的任一解. 则我们可选取 $\epsilon>0$ 充分小使得对所有的 $t\geqslant 0$ 都有 $|u_2(t)|\leqslant|u_1(t)|$.

证明　变分方程可以写成

$$u_1'=\lambda u_1+f_{11}u_1+f_{12}u_2$$

$$u_2'=-\mu u_2+f_{21}u_1+f_{22}u_2,$$

通过选取更小的 ϵ, 可以使得 $|f_{ij}(t)|$ 要多小有多小. 任给该方程的解 $(u_1(t),u_2(t))$, 记 $\eta=\eta(t)$ 为这个解的斜率 $u_2(t)/u_1(t)$. 利用变分方程以及函数的商的求导公式, 可见 η 满足

$$\eta'=\alpha(t)-(\mu+\lambda+\beta(t))\eta+\gamma(t)\eta^2,$$

其中 $\alpha(t)=f_{21}(t),\beta(t)=f_{11}(t)-f_{22}(t),\gamma(t)=-f_{12}(t)$. 特别地, α,β,γ 都很小. 想到 $\mu+\lambda>0$, 于是

$$-(\mu+\lambda+\beta(t))<0.$$

由此可得, 如果 $\eta(t)>0$, 则存在常数 $a_1,a_2,0<a_1<a_2$ 使得

$$-a_2\eta+\alpha(t)\leqslant\eta'\leqslant-a_1\eta+\alpha(t).$$

现在, 如果取 ϵ 充分小, 且令 $u_1(0)=u_2(0)$, 从而 $\eta(0)=1$, 则有 $\eta'(0)<0$, 从而在开始时 $\eta(t)$ 是递减的. 类似地讨论可得, 如果 $\eta(0)=-1$, 则 $\eta'(0)>0$, 从而 $\eta(t)$ 递增. 特别地, 任何在开始满足 $|\eta(0)|<1$ 的解将永远不会满足 $\eta(t)=\pm1$. 由于我们的解满足 $|\eta(0)|\leqslant 1$, 从而这个解将会陷入区域 $|\eta(t)|<1$, 于是我们就有, 对所有的 $t>0$, $|u_2(t)|\leqslant|u_1(t)|$. 这就证明了引理. ∎

为了完成定理的证明, 假设有第二个永远位于 C_M 中的解 $\boldsymbol{Y}(t)$. 于是当 $t\to\infty$ 时, $\boldsymbol{Y}(t)\to 0$. 假设 $\boldsymbol{Y}(0)=(x_0',\epsilon),x_0'\neq x_0$. 可见, 当初始条件 (x,ϵ) 满足 $x_0<x<x_0'$ 时, 从 (x,ϵ) 出发的解也趋于原点, 于是我们不妨假设 x_0' 足够接近 x_0. 我们将用变分方程的一个解来逼近 $Y(t)-X(t)$. 记 $\boldsymbol{U}(t)$ 为变分方程

$$\boldsymbol{U}'=\boldsymbol{DF}_{\boldsymbol{X}(t)}\boldsymbol{U}$$

满足 $u_1(0)=x_0'-x_0,u_2(0)=0$ 的解. 将上面的引理应用到 $\boldsymbol{U}(t)$ 可得, 对所有的 t, $u_2(t)\leqslant u_1(t)$.

我们有

$$\begin{aligned}|u_1'| &= |\lambda u_1 + f_{11}(t)u_1 + f_{12}(t)u_2| \\ &\geqslant \lambda|u_1| - (|f_{11}(t)| + |f_{12}(t)|)|u_1|.\end{aligned}$$

但由于当 $\epsilon \to 0$ 时 $|f_{11} + f_{12}|$ 趋于 0, 于是我们有, 对于某一 $K > 0$,

$$|u_1'| \geqslant K|u_1|.$$

于是, 对于 $t \geqslant 0$, $|u_1(t)| \geqslant C\mathrm{e}^{Kt}$. 又根据假设 $\boldsymbol{Y}(t)$ 趋于 0. 但是, 如果 $u_1(0) \neq 0$, 则 $\boldsymbol{Y}(t)$ 和 $\boldsymbol{U}(t)$ 之间的距离必定要变得很大. 这个矛盾就证明了 $\boldsymbol{Y}(t) = \boldsymbol{X}(t)$.

最后, 我们断言, $\boldsymbol{X}(t)$ 是切于 y 轴趋于原点的. 事实上, 不管 M 多大, 我们总可选取 ϵ 充分小, 使得 $\boldsymbol{X}(t)$ 位于 C_M 中, 于是, 当 $\boldsymbol{X}(t)$ 逼近 $(0,0)$ 时, $\boldsymbol{X}(t)$ 的斜率趋于 $\pm\infty$. 这就证明了我们的断言. ■

我们简单地讨论一下高维的鞍点来结束本节. 假设 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$, $\boldsymbol{X} \in \mathbb{R}^n$. 假设 $\boldsymbol{X}_0$ 为一个平衡解, 其相应的线性化系统具有 k 个实部为负的特征值和 $n-k$ 个实部为正的特征值. 则局部稳定集和局部不稳定集一般不再是曲线, 而分别是 k 维和 $n-k$ 维的 "子流形". 我们不打算在此讨论流形理论, 而只是简单地指出, 这意味着, 在经过一个线性坐标变换后, 原点附近的局部稳定集由一个 C^∞ 函数 $g: B_r \to \mathbb{R}^{n-k}$ 的图像所给出, 其中 $g(0) = 0$, 且 g 的所有的偏导数在原点都是零. 这里, B_r 是 $\mathbb{R}^k$ 中中心在原点半径为 r 的圆盘. 局部不稳定集是一个 $n-k$ 维圆盘上的类似的图像. 这两个图像在平衡点 $\boldsymbol{X}_0$ 处分别切于它的稳定子空间和不稳定子空间. 于是它们只在 $\boldsymbol{X}_0$ 处相交.

例 考虑系统

$$\begin{aligned}x' &= -x \\ y' &= -y \\ z' &= z + x^2 + y^2.\end{aligned}$$

其在原点的线性化系统的特征值为 1 和 -1(二重). 坐标变换

$$\begin{aligned}u &= x \\ v &= y \\ w &= z + \frac{1}{3}(x^2 + y^2)\end{aligned}$$

将这个非线性系统化成下面的线性系统:

$$\begin{aligned}u' &= -u \\ v' &= -v \\ w' &= w.\end{aligned}$$

平面 $w=0$ 为这个线性系统的稳定平面. 上面的坐标变换将这个平面变换成曲面:

$$z=-\frac{1}{3}(x^2+y^2),$$

这是 $\mathbb{R}^3$ 中过原点的一个抛物面, 而且开口向下. 这个曲面上的所有解都趋于原点; 我们称它为这个非线性系统的**稳定曲面** (见图 8.5). ■

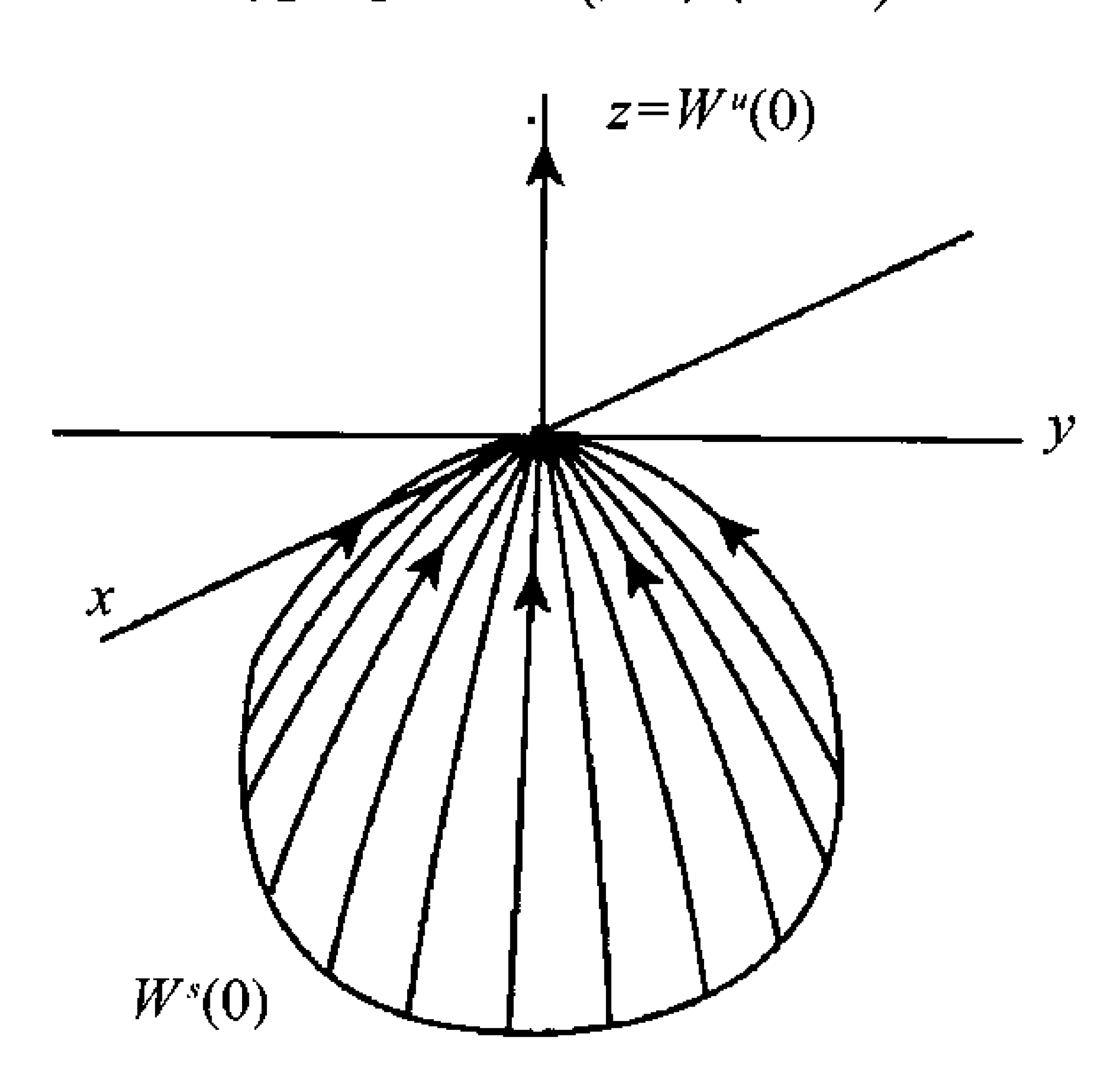


图 8.5　$x'=-x, y'=-y, z'=z+x^2+y^2$ 的相图

8.4 稳　定　性

在常微分方程及其应用中, 平衡点的研究扮演了重要的角色. 然而, 为了保证一个平衡点有物理意义 (此处以及本书中其他一些地方, 我们都是在一种广义的意义下使用**物理**这个词, 在某些场合, 物理可能要被替换成**生物**、**化学**, 甚至是**经济**), 它必须满足一定的稳定性假设.

一个平衡点称为**稳定**的, 如果附近的解在将来都仍然逗留在附近. 在动力系统的应用中, 人们常常不能给出位置的准确值, 而只是一个近似值, 因而为了平衡点具有物理意义, 它必须是稳定的.

更准确地说, 假设 $\boldsymbol{X}^*\in\mathbb{R}^n$ 为微分方程

$$\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$$

的一个平衡点. 则 $\boldsymbol{X}^*$ 为一个**稳定的**平衡点, 如果对 $\boldsymbol{X}^*$ 在 $\mathbb{R}^n$ 中的任一邻域 $\mathcal{O}$, 都存在 $\boldsymbol{X}^*$ 在 $\mathcal{O}$ 中的邻域 $\mathcal{O}_1$, 使得任何满足初始条件 $\boldsymbol{X}(0)=\boldsymbol{X}_0\in\mathcal{O}_1$ 的解 $\boldsymbol{X}(t)$ 对所有的 $t>0$, 都始终逗留在 $\mathcal{O}$ 中.

另一种形式不同的稳定性是**渐近稳定性**. 如果上面选取的 $\mathcal{O}_1$ 除了能保证稳定性外, 还满足

$$\lim_{t\to\infty}\boldsymbol{X}(t)=\boldsymbol{X}^*,$$

则我们称 $\boldsymbol{X}^*$ 是**渐近稳定的**.

如果平衡点 $\boldsymbol{X}^*$ 不是稳定的, 则称为**不稳定**的. 这意味着存在 $\boldsymbol{X}^*$ 的一个邻域 $\mathcal{O}$ 使得对于 $\boldsymbol{X}^*$ 在 $\mathcal{O}$ 中的**每一个**邻域 $\mathcal{O}_1$, 都至少存在一个从 $\boldsymbol{X}(0) \in \mathcal{O}_1$ 出发的解 $\boldsymbol{X}(t)$ 并不是对于所有的 $t > 0$ 都整个地位于 $\mathcal{O}$ 中.

汇点是渐近稳定的, 因而也是稳定的. 源点和鞍点都是不稳定平衡点的例子. 对于 $\mathbb{R}^2$ 中的线性方程 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, 如果 $\boldsymbol{A}$ 具有纯虚的特征值, 则原点就是一个稳定但非渐近稳定的平衡点. 这个例子在应用中的重要性是很有限的 (著名的调和振子除外), 这么说的理由是, 在 8.1 节中, 我们已经看到, 任何微小的非线性扰动都可能破坏它的这一性质; 甚至是一个小的线性扰动都能够将一个中心变成一个汇点或源点.

于是, 如果系统在一个平衡点处的线性化系统是双曲的, 则我们可以立刻确定该平衡点的稳定性. 遗憾的是, 应用中出现的许多重要的平衡点都是非双曲的. 如果有一个技巧能够确定所有情形的平衡点的稳定性, 那就太好了. 令人再次遗憾的是, 确定稳定性, 没有这样万能的方法, 而只能通过找出系统的所有的解, 而这即使可能, 也往往是很困难的. 在下一章, 我们将给出几种技巧使得在一些特殊情形, 我们可以确定稳定性.

8.5 分　岔

在这一节, 我们将描述非线性系统中出现的一些简单的分岔例子. 我们考虑一族系统

$$X' = F_a(X),$$

其中 a 为实参数. 我们假设 F_a 以 C^∞ 的方式依赖于 a. 当 a 变化时, 如果系统的解的结构发生了 "显著的" 改变, 就称出现了一个分岔. 最简单的分岔就是当 a 变化时, 平衡解的数目发生了变化.

回忆一下我们在第 1 章中遇到的一阶方程 $x' = f_a(x)$ 的初等分岔. 当 x_0 为一平衡点时, 我们有 $f_a(x_0) = 0$. 如果 $f_a'(x_0) \neq 0$, 则 a 的微小变化将不会改变 x_0 附近的局部结构: 即, 当 ϵ 很小时, 微分方程

$$x' = f_{a+\epsilon}(x)$$

有一个关于 ϵ 连续变化的平衡点 $x_0(\epsilon)$. 看一看 x_0 附近的 $f_{a+\epsilon}(x)$ 的 (递增或递减的) 图像就可知道为什么是这样的. 更严格地说, 这是隐函数定理的一个直接推论 (见本章的习题 3). 由此可见, 一阶方程只有在非双曲情形才会出现分岔, 此时有 $f_a'(x_0) = 0$.

例 考虑一阶方程

$$x' = f_a(x) = x^2 + a.$$

当 $a = 0$ 时, $x = 0$ 是它的唯一的平衡点. 注意, $f_0'(0) = 0$, 而 $f_0''(0) \neq 0$. 当 $a > 0$ 时, 对所有的 x, 都有 $f_a(x) > 0$, 因而方程没有平衡点; 但当 $a < 0$ 时, 该方程有一对平衡点. 于是, 当参数经过 $a = 0$ 时, 出现了分岔. ■

这类分岔称为**鞍结分岔**(稍后我们将看到这个分岔中的 "鞍点"). 在鞍结分岔中, 存在包含分岔值 a_0 的一个区间以及另一个在 x 轴上的区间 I, 使得该微分方程

(1) 当 $a < a_0$ 时在 I 中有两个平衡点;

(2) 当 $a = a_0$ 时在 I 中有一个平衡点;

(3) 当 $a > a_0$ 时在 I 中没有平衡点.

当然, 分岔也可以 "按相反的方向" 出现, 所谓相反的方向是指当 $a < a_0$ 时没有平衡点. 实际上, 上面的例子就是一阶方程的典型分岔.

定理 (**鞍结分岔**) 假设一阶微分方程 $x' = f_a(x)$ 满足

(1) $f_{a_0}(x_0) = 0$; (2) $f_{a_0}'(x_0) = 0$; (3) $f_{a_0}''(x_0) \neq 0$; (4) $\dfrac{\partial f_{a_0}}{\partial a}(x_0) \neq 0$.

则该微分方程在 $a = a_0$ 处出现一个鞍结分岔.

证明 记 $G(x, a) = f_a(x)$. 我们有 $G(x_0, a_0) = 0$, 以及

$$\frac{\partial G}{\partial a}(x_0, a_0) = \frac{\partial f_{a_0}}{\partial a}(x_0) \neq 0,$$

于是, 由隐函数定理可得, 存在光滑函数 $a = a(x)$ 使得 $G(x, a(x)) = 0$. 特别地, 如果 x^* 在 $a(x)$ 的定义域中, 则 $f_{a(x^*)}(x^*) = 0$, 从而它就是方程 $x' = f_{a(x^*)}(x)$ 的一个平衡点. 在 $G(x, a(x)) = 0$ 中对 x 求导, 我们求得

$$a'(x) = \frac{-\partial G/\partial x}{\partial G/\partial a}.$$

现在, 根据假设, 我们有 $(\partial G/\partial x)(x_0, a_0) = f_{a_0}'(x_0) = 0$, 而 $(\partial G/\partial a)(x_0, a_0) \neq 0$, 从而 $a'(x_0) = 0$. 再求一次导数可得

$$a''(x) = \frac{-\dfrac{\partial^2 G}{\partial x^2}\dfrac{\partial G}{\partial a} + \dfrac{\partial G}{\partial x}\dfrac{\partial^2 G}{\partial x \partial a}}{\left(\dfrac{\partial G}{\partial a}\right)^2}.$$

由于 $(\partial G/\partial x)(x_0, a_0) = 0$, 我们有

$$a''(x_0) = \frac{-\dfrac{\partial^2 G}{\partial x^2}(x_0, a_0)}{\dfrac{\partial G}{\partial a}(x_0, a_0)} \neq 0,$$

这是因为 $(\partial^2 G/\partial x^2)(x_0,a_0)=f''_{a_0}(x_0)\neq 0$. 这意味着, $a=a(x)$ 的图像要么是上凹的, 要么是下凹的, 从而当 a 在 a_0 的一边取值时, 我们在 x_0 附近有两个平衡点, 而在另一边取值时没有平衡点. ■

前面我们说过, 这种鞍结分岔在一阶方程关于平衡点的分岔中是"典型的". 这种说法的理由如下: 如果 $x'=f_a(x)$ 在 $a=a_0$ 时有一个分岔, 则我们必定有

(1) $f_{a_0}(x_0)=0$ 和 (2)$f'_{a_0}(x_0)=0$.

一般地 (在 5.6 节的意义下), 在 (x_0,a_0) 处的下一个高阶导数将是非零的, 即, 在这样的一个分岔点处, 下面的两条

(3) $f''_{a_0}(x_0)\neq 0$ 和 (4) $\dfrac{\partial f_a}{\partial a}(x_0,a_0)\neq 0$

将是典型的. 而这些正是保证出现鞍结分岔的条件.

回忆一下, 所谓 $x'=f_a(x)$ 的分岔图就是将不同参数 a 所对应的方程的相直线画在一起形成的图. 图 8.6 所示的是一个典型的鞍结分岔的分岔图. (其中箭头的方向和平衡点的曲线可能会不同.)

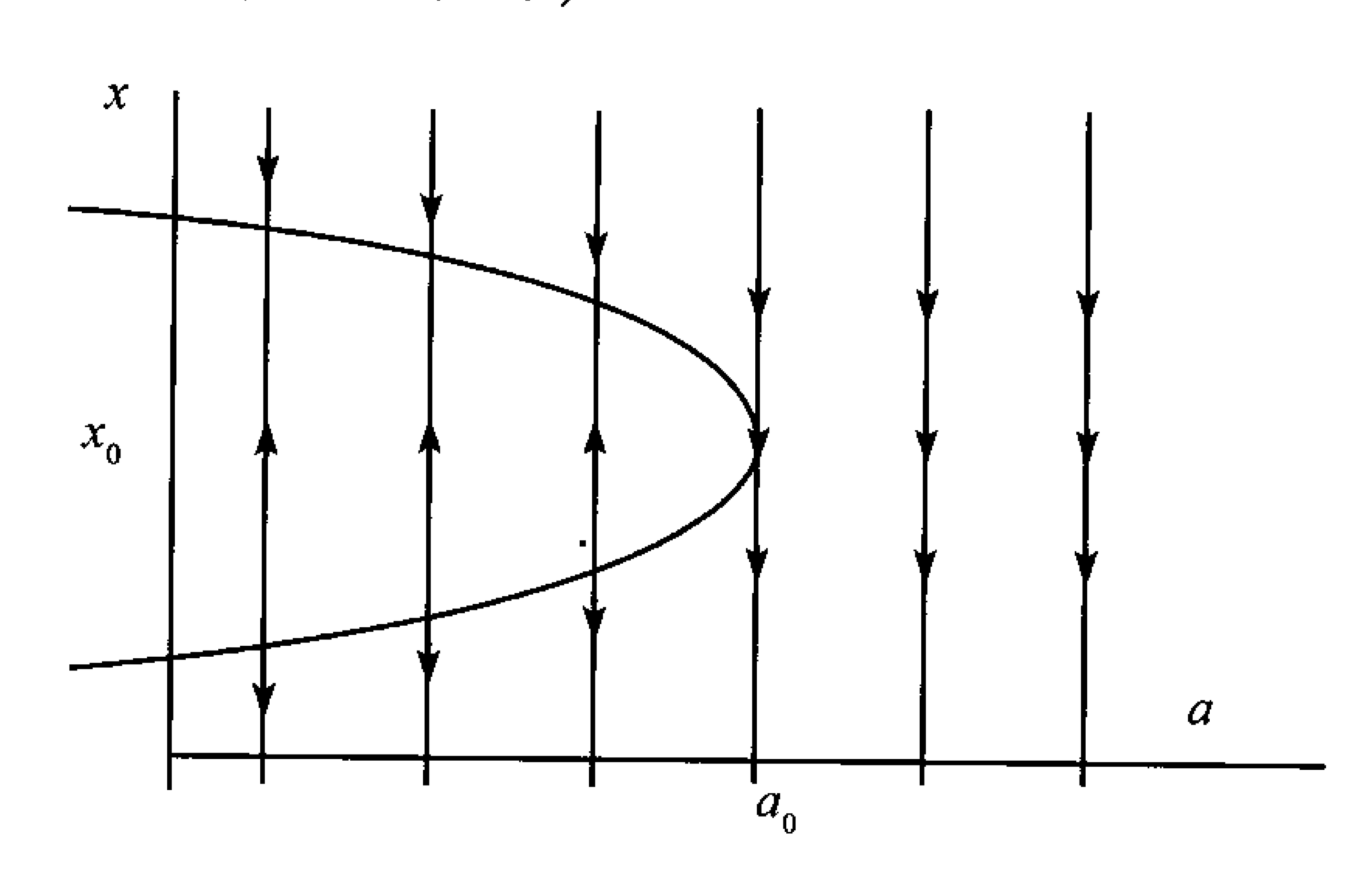


图 8.6 一个鞍结分岔的分岔图

例 **(叉分岔)**考虑

$$x'=x^3-ax.$$

当 $a>0$ 时, 该方程有 3 个平衡点, 分别是 $x=0$ 以及 $x=\pm\sqrt{a}$. 当 $a\leqslant 0$ 时, $x=0$ 是唯一的平衡点. 其分岔图见图 8.7, 从中可以看出为什么这样称呼这种分岔. ■

现在我们来看几个高维的分岔. 平面中的鞍结分岔类似于它相应的一维情形, 只不过现在可以看出"鞍点"在哪里.

例 考虑系统

$$x'=x^2+a$$
$$y'=-y.$$

当 $a=0$ 时, 这是 8.1 节中曾经考虑过的一个系统. 此时, 方程在原点处有唯一的平衡点, 而且相应的线性化系统有一个零特征值.

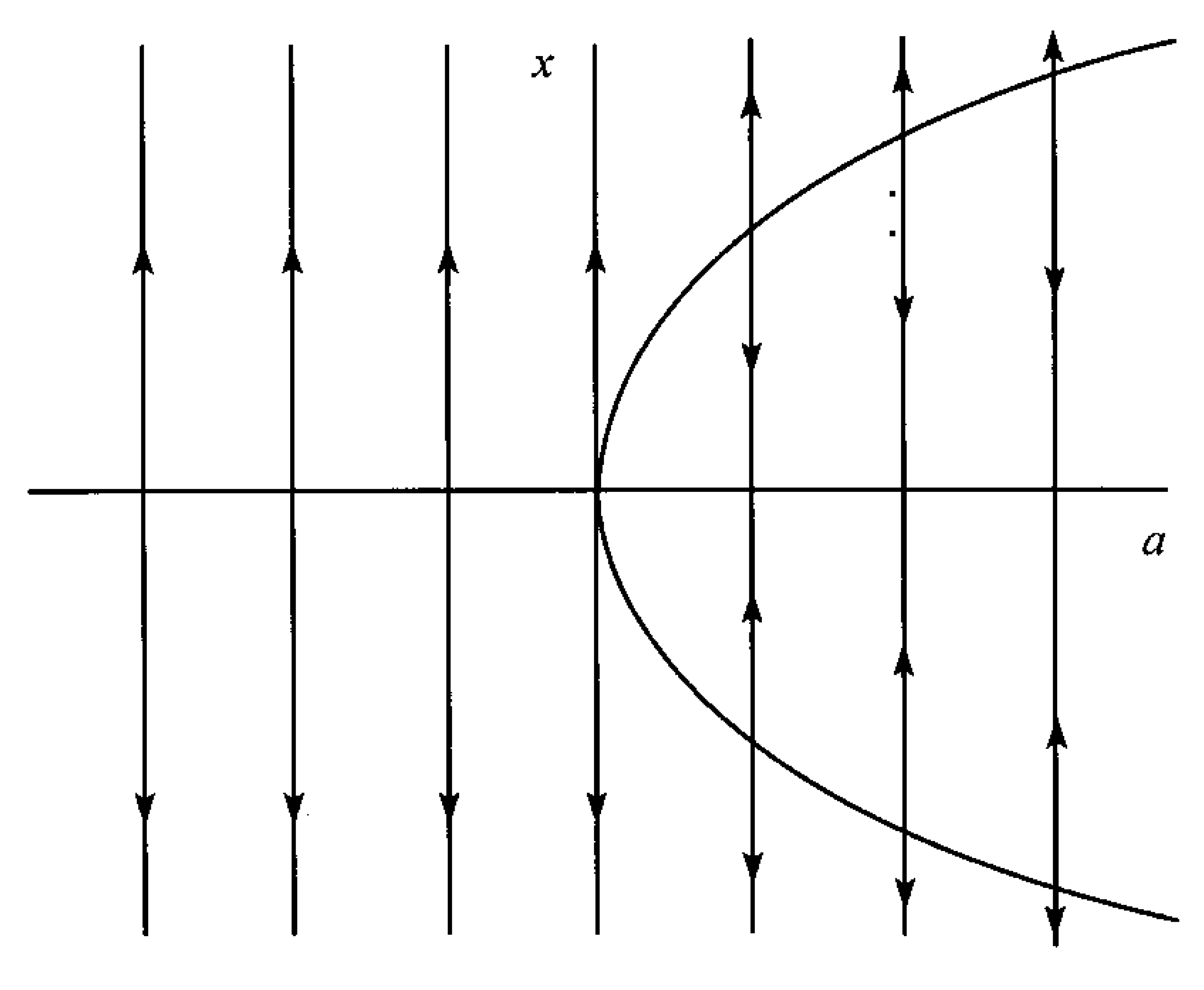


图 8.7　一个叉分岔的分岔图

当 a 经过 $a=0$ 时, 鞍结分岔就出现了. 当 $a>0$ 时, 我们有 $x'>0$, 因而所有的解都向右边运动; 此时, 平衡点不见了. 当 $a<0$ 时, 我们有一对平衡点, 它们是 $(\pm\sqrt{-a},0)$. 相应的线性化系统是

$$X' = \begin{pmatrix} 2x & 0 \\ 0 & -1 \end{pmatrix} X.$$

从而 $(-\sqrt{-a},0)$ 是一个汇点, $(\sqrt{-a},0)$ 是一个鞍点. 注意, 在直线 $x=\pm\sqrt{-a}$ 上, $x'=0$, 因而这两条线上的所有的解都始终逗留在这两条线上. 又由于 $y'=-y$, 这些解都笔直地趋于相应的平衡点. 该分岔的略图见图 8.8. ■

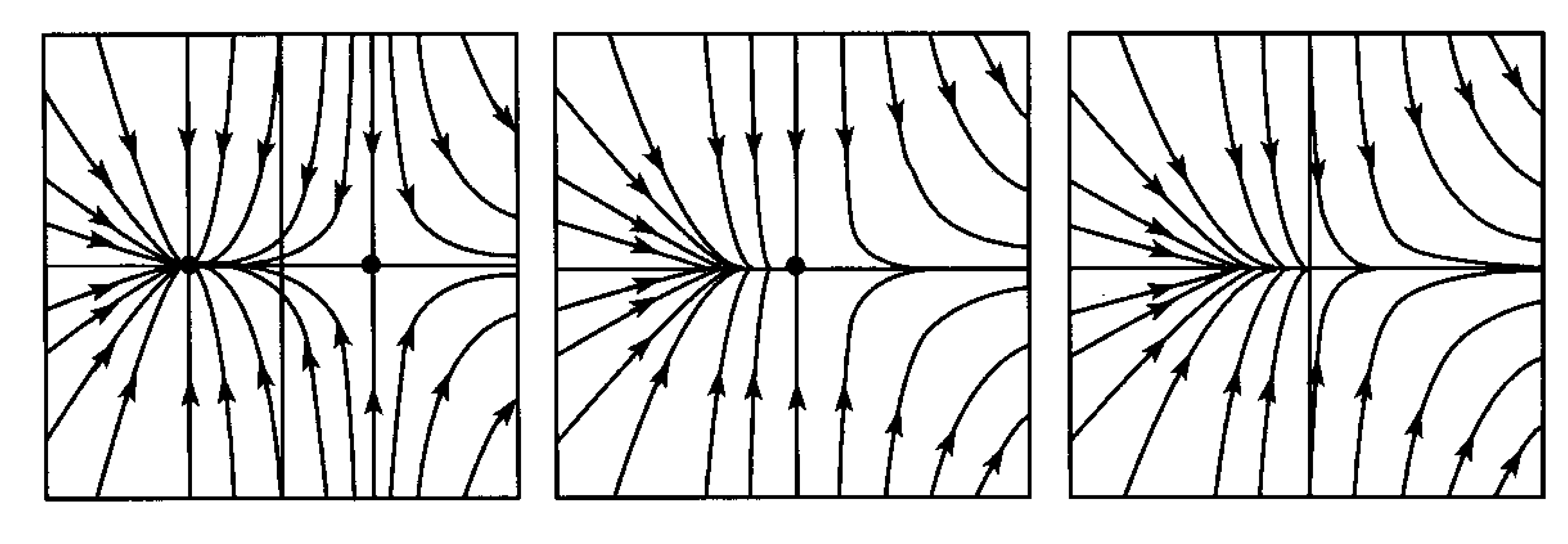


图 8.8　鞍结分岔, 其中从左到右, 分别对应于 $a<0, a=0, a>0$

下面的例子表明, 鞍结分岔可能对解的大范围行为产生影响.

例　考虑由极坐标给出的系统

$$r' = r - r^3$$
$$\theta' = \sin^2\theta + a,$$

这里的 a 也是一个参数. 由于当 $r=0$ 时, $r'=0$, 所以原点总是一个平衡点. 由于当 $a>0$ 时, $\theta'>0$, 从而此时方程没有其他的平衡点. 当 $a=0$ 时, 将多出 2 个平

衡点, 它们分别位于 $(r,\theta)=(1,0)$ 和 $(r,\theta)=(1,\pi)$. 当 $-1<a<0$ 时, 在圆周 $r=1$ 上有 4 个平衡点, 它们分别对应于方程

$$\sin^2\theta=-a$$

的 4 个根. 将这 4 个根记为 $\theta_\pm,\theta_\pm+\pi$, 这里我们假设 $0<\theta_+<\pi/2,-\pi/2<\theta_-<0$.

由于 θ' 只依赖于 θ, 而与 r 无关, 因而该系统的流将过原点的射线 $\theta=c$(c 为常数) 带到另外的射线. 另外, 由于在单位圆周上有 $r'=0$, 因而从这个圆周上出发的解都始终逗留在该圆周上, 或换一种说法, 此时我们称单位圆周是不变的. 又由于当 $0<r<1$ 时, $r'>0$, 以及当 $r>1$ 时, $r'<0$, 因而所有其他的非零解都趋向圆周.

现在考虑 $a=0$ 的情形. 此时, x 轴是不变的, 且该直线上的所有非零解都趋于在 $x=\pm1$ 处的平衡点. 在上半平面, 我们有 $\theta'>0$, 因而这个区域中的所有其他解都绕原点反时针转动, 且趋于 $x=-1$; 相应的 θ 坐标递增地趋于 $\theta=\pi$, 而 r 坐标则单调趋于 $r=1$. 因为有 x 轴这个障碍, 任何解绕原点转动的角度都不会超过 π. 在下半平面, 解的行为是对称的.

当 $a>0$ 时, 出现了两件事. 首先, 在 $x=\pm1$ 处的平衡点消失了, 而且此时, 处处都有 $\theta'>0$. 由于移走了 x 轴这个障碍, 从而所有的解都可以永远绕原点转动下去. 其次, 在圆周 $r=1$ 上, 我们得到一个周期解, 它将吸引所有的非零解.

这些巨大的变化都是因为在 $a=0$ 处出现了一对鞍结分岔. 事实上, 当 $-1<a<0$ 时, 在单位圆周上, 我们有 2 对平衡点. 射线 $\theta=\theta_\pm$ 和 $\theta=\theta_\pm+\pi$ 都是不变的, 这些射线上的解都趋于单位圆周上的平衡点. 考虑半平面 $\theta_-<\theta<\theta_-+\pi$. 当 θ 取值于区间 $\theta_-<\theta<\theta_+$ 时, 我们有 $\theta'<0$, 而当 θ 取值于区间 $\theta_+<\theta<\theta_-+\pi$ 时, 我们有 $\theta'>0$. 在余下的半平面上, 解的行为是对称的. 由此可得, 所有不在射线 $\theta=\theta_+$ 或 $\theta=\theta_++\pi$ 上的解要么趋于平衡点 $r=1,\theta=\theta_-$, 要么趋于平衡点 $r=1,\theta=\theta_-+\pi$. 这 2 个平衡点都是汇点, 而单位圆周上另外 2 个平衡点则都是鞍点. 2 个鞍点的稳定曲线分别位于射线 $\theta=\theta_+$ 和 $\theta=\theta_++\pi$ 上, 而不稳定曲线则分别是单位圆周去掉 2 个汇点后的单位圆 (见图 8.9). ■

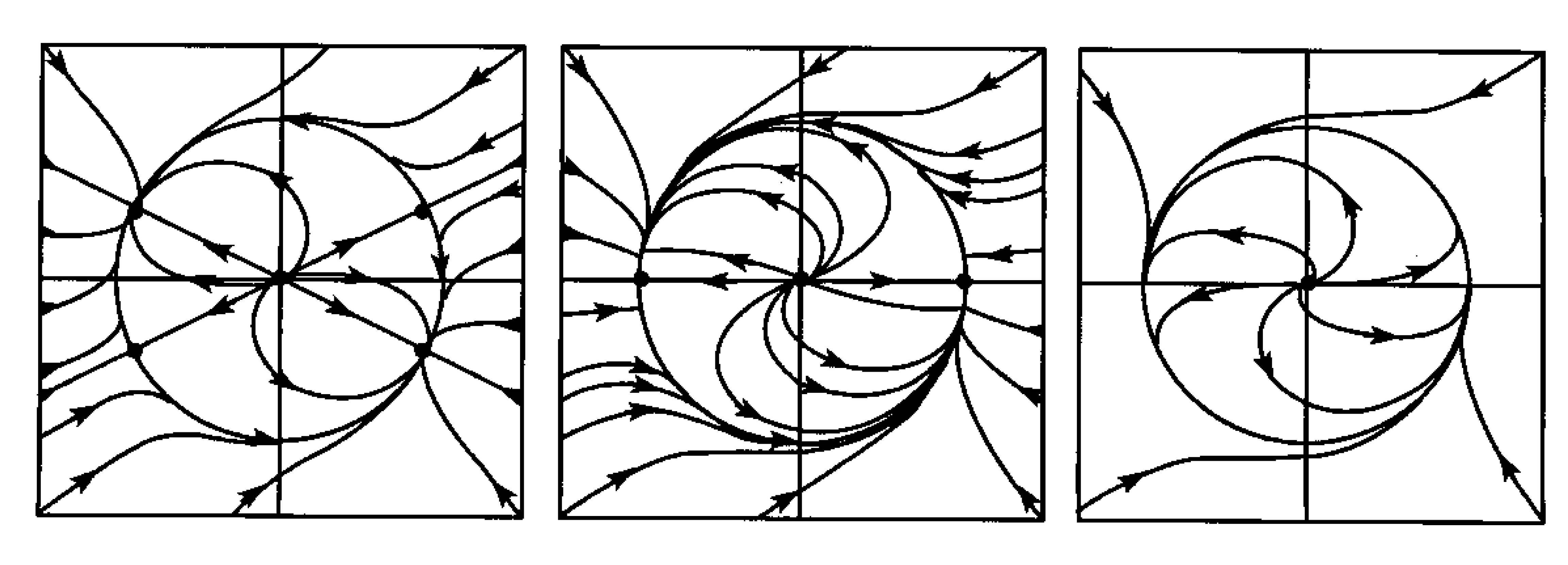


图 8.9 鞍结分岔的大范围效应, 其中从左到右, 分别对应于 $a<0,a=0,a>0$

前面的例子都是描述当线性系统具有零特征值时所出现的分岔的特点.

例 (**霍普夫分岔**) 考虑系统

$$x' = ax - y - x(x^2 + y^2)$$

$$y' = x + ay - y(x^2 + y^2).$$

该系统在原点有一个平衡点, 其线性化系统为

$$X' = \begin{pmatrix} a & -1 \\ 1 & a \end{pmatrix} X,$$

其特征值为 $a \pm \mathrm{i}$, 因而我们期望在 $a = 0$ 时出现分岔.

为了看出当 a 经过 0 时发生的现象, 我们变换到极坐标下考虑. 此时系统变成

$$r' = ar - r^3$$

$$\theta' = 1.$$

由于 $\theta' \neq 0$, 可见原点为该系统的唯一平衡点. 当 $a < 0$ 时, 由于对所有的 $r > 0$, $ar - r^3 < 0$, 故原点是一个汇点. 从而, 此时所有的解都趋于原点. 当 $a > 0$ 时, 平衡点变成为一个源点. 但是, 还发生了些什么? 当 $a > 0$ 时, 如果 $r = \sqrt{a}$, 则 $r' = 0$, 从而半径为 $\sqrt{a}$ 的圆周是一个周期为 2π 的周期解. 此时, 我们还有, 如果 $0 < r < \sqrt{a}$, 则 $r' > 0$, 而如果 $r > \sqrt{a}$ 则 $r' < 0$, 从而所有的非零解当 $t \to \infty$ 时都盘旋地趋向这个圆周解.

这种类型的分岔称为**霍普夫分岔**. 于是, 在一个霍普夫分岔中, 不出现新的平衡点, 而是当 a 经过分岔值时在平衡点产生了一个周期解 (见图 8.10). ■

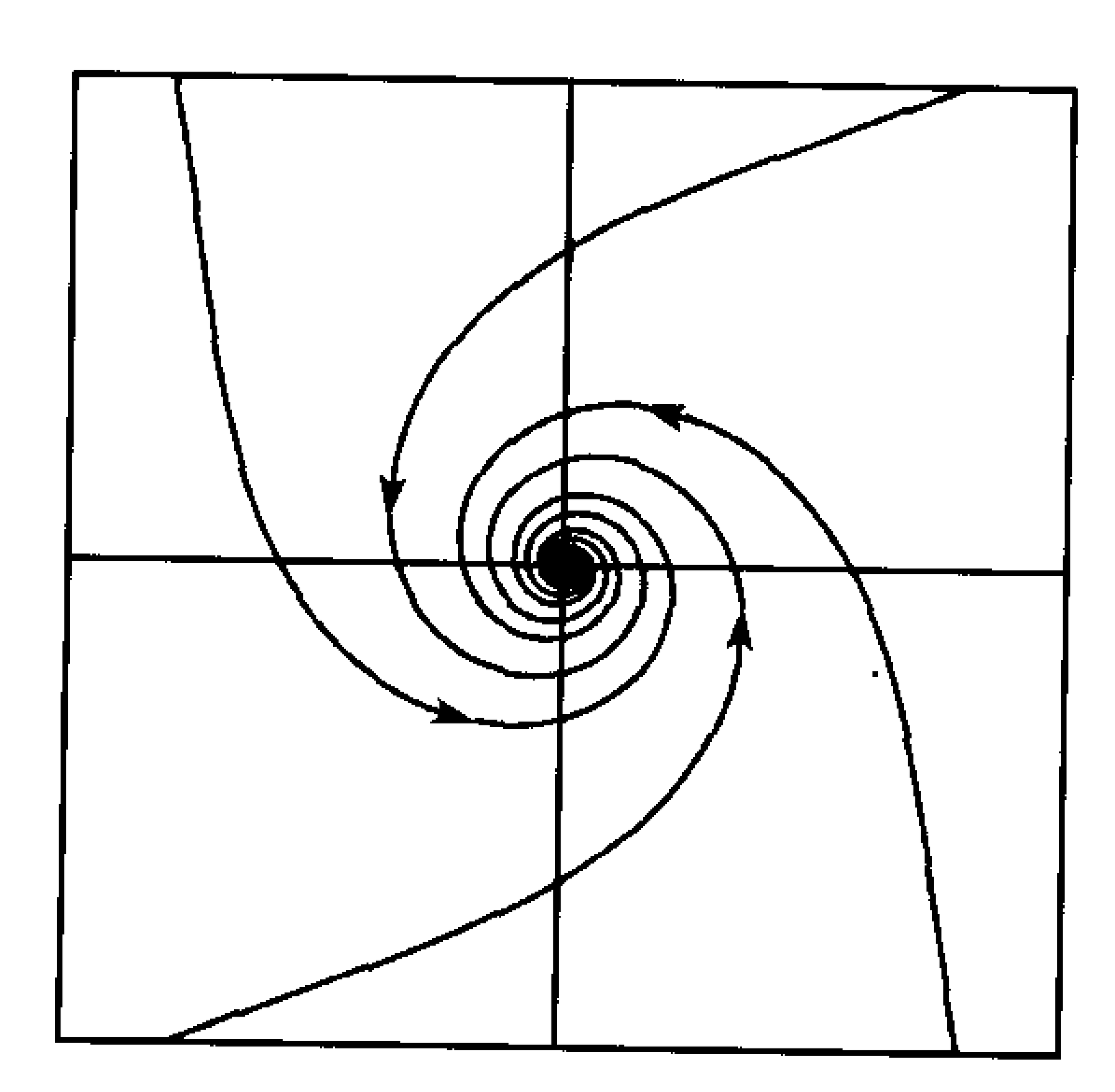
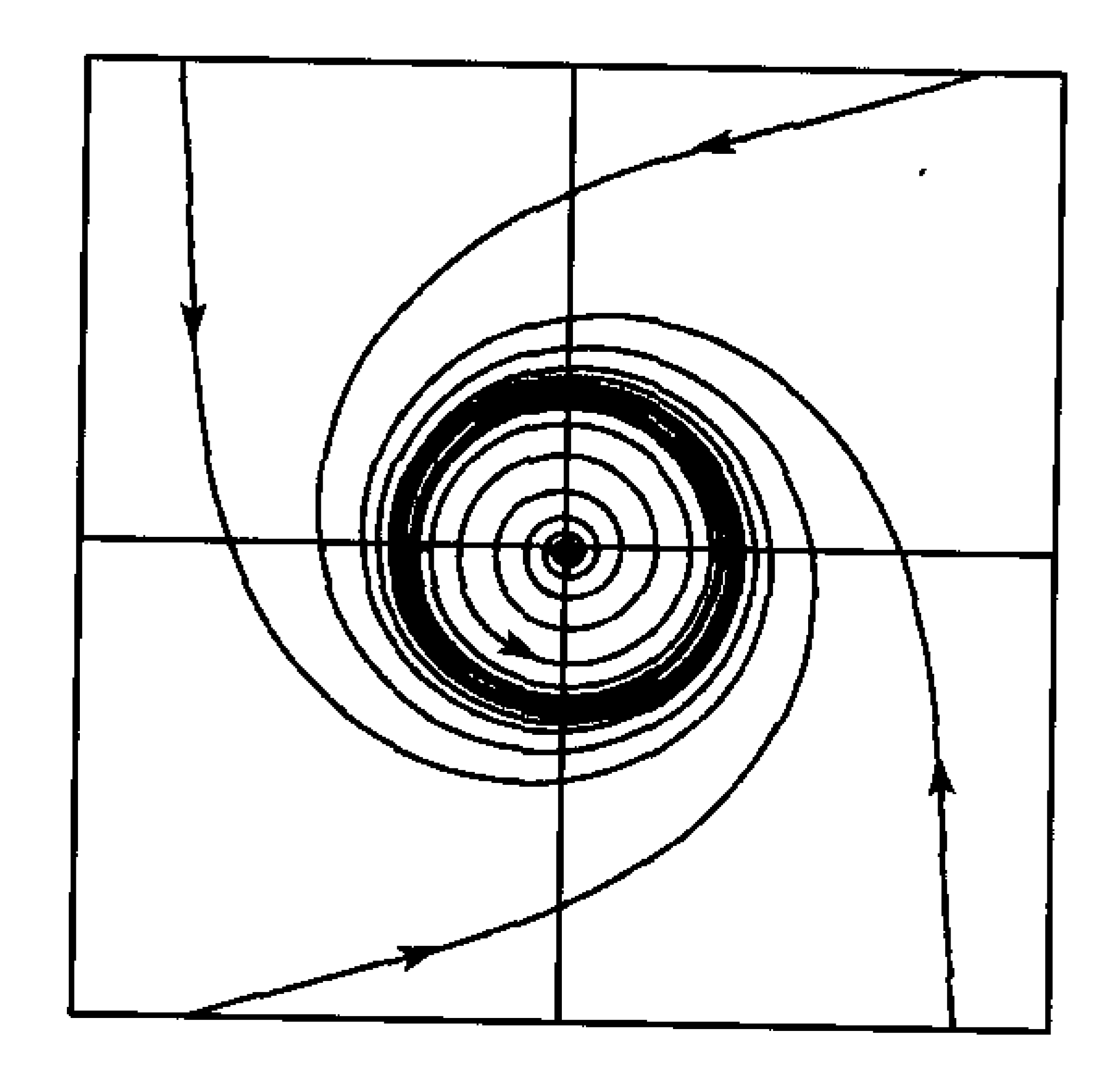



图 8.10 霍普夫分岔在 $a < 0$ 和 $a > 0$ 时的情形

8.6 探索：复向量场

在这个探索中，我们研究复平面上形如 $z' = F(z)$ 的微分方程系统的行为. 在本节中，z 都代表复数 $z = x + \mathrm{i}y$，而 $F(z)$ 则是一个复系数的多项式. 微分方程的解将用复平面上的曲线 $z(t) = x(t) + \mathrm{i}y(t)$ 表示.

为了完全理解下面的内容，你们必须熟悉诸如指数、正弦、余弦等复函数，以及求复平方根的过程. 理论上，你们还必须对复分析有所掌握. 然而，所有那些对实变量函数求积分的通常技巧在对 z 积分时也都同样成立. 你们不必证明这点，这是因为当完成积分之后，你们总是可以验算所求解的正确性.

(1) 求解方程 $z' = az$，其中 a 为复数. 对于这些微分方程，你们在原点处能发现什么类型的平衡点？

(2) 求解下列的复微分方程并简单地作出相图.

(a) $z' = z^2$; (b) $z' = z^2 - 1$; (c) $z' = z^2 + 1$.

(3) 对于一个复多项式 $F(z)$，可以像定义实导数一样定义它的复导数

$$F'(z_0) = \lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0},$$

只不过，现在求极限要沿 $\mathbb{C}$ 中任一条过 z_0 的光滑曲线来进行. 这个极限必定存在，且对于每条这样的曲线都得到同一个 (复) 数. 对于多项式，这一点很容易得到验证. 现在记

$$F(x + \mathrm{i}y) = u(x, y) + \mathrm{i}v(x, y).$$

先沿水平线 $z_0 + t$ 求极限，然后沿竖直线 $z_0 + \mathrm{i}t$ 求极限，分别计算 $F'(z_0)$，结果用 u, v 的导数表示. 利用这两个结果推出，如果导数存在；则在复平面上的每个点处，我们都一定有

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{和} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

这两个方程称为**柯西-黎曼方程.**

(4) 利用上面的观察确定复向量场的平衡点的所有可能类型.

(5) 求解方程

$$z' = (z - z_0)(z - z_1),$$

其中 $z_0, z_1 \in \mathbb{C}$，且 $z_0 \neq z_1$. 对于 z_0, z_1 的不同取值会出现哪些类型的平衡点？

(6) 找一个变量的非线性变换，使其将前面的系统变为 $w' = \alpha w (\alpha \in \mathbb{C})$.

提示：由于原始系统有两个平衡点，而线性系统只有一个平衡点，故变量的变换必定将其中的一个平衡点变为 ∞.

(7) 将形如

$$z' = z^2 + az + b$$

的复二次系统分类, 其中 $a, b \in \mathbb{C}$.

(8) 考虑方程

$$z' = z^3 + az,$$

其中 $a \in \mathbb{C}$. 先用计算机描述这些系统的相图, 然后对这些系统证明尽可能多的结果, 并将它们相应于 a 分类.

(9) 自己选一族 (非平凡的) 依赖于参数 $a \in \mathbb{C}$ 的复函数并且对每个 a, 完整地分析它的相图. 可以考虑的简单族包括 $a \exp z, a \sin z$ 或 $(z^2 + a)(z^2 - a)$.

习 题

1. 对下面的每个非线性系统
   (a) 找出所有的平衡点并描述相应的线性化系统的行为.
   (b) 描述非线性系统的相图.
   (c) 线性化系统是否准确地描述了平衡点附近的局部行为?
   (i) $x' = \sin x,\ y' = \cos y$;
   (ii) $x' = x(x^2 + y^2),\ y' = y(x^2 + y^2)$;
   (iii) $x' = x + y^2,\ y' = 2y$;
   (iv) $x' = y^2,\ y' = y$;
   (v) $x' = x^2,\ y' = y^2$.
2. 找出一个大范围的坐标变换, 使得它将下面的系统线性化.

$$x' = x + y^2$$
$$y' = -y$$
$$z' = -z + y^2.$$

3. 考虑一阶微分方程

$$x' = f_a(x),$$

假设它满足 $f_a(x_0) = 0, f_a'(x_0) \neq 0$. 证明微分方程

$$x' = f_{a+\epsilon}(x)$$

在 ϵ 充分小时, 有一个平衡点 $x_0(\epsilon)$, 而且 $\epsilon \to x_0(\epsilon)$ 是一个满足 $x_0(0) = x_0$ 的光滑函数.

4. 找出关于 $f_a(x)$ 的导数的一般的条件, 使得方程

$$x' = f_a(x)$$

在 $a = a_0$ 处发生叉分岔. 证明你找到的条件可以导致这样的分岔.

5. 考虑系统

$$\begin{aligned} x' &= x^2 + y \\ y' &= x - y + a, \end{aligned}$$

其中 a 是一个参数.

(a) 找出所有的平衡点并计算每个平衡点处的线性化方程.

(b) 描述在每个平衡点处的线性化系统的行为.

(c) 描述所有出现的分岔.

6. 举例说明, 存在一族微分方程 $x' = f_a(x)$ 满足：当 $a < 0$ 时没有平衡点；当 $a = 0$ 时只有一个平衡点；当 $a > 0$ 时有 4 个平衡点. 简略地作出这族方程的分岔图.

7. 讨论方程

$$\begin{aligned} r' &= r - r^3 \\ \theta' &= \sin^2\theta + a \end{aligned}$$

在分岔值 $a = -1$ 处的解的局部行为和大范围行为.

8. 讨论方程

$$\begin{aligned} r' &= r - r^2 \\ \theta' &= \sin^2(\theta/2) + a \end{aligned}$$

在所有分岔值处的解的局部行为和大范围行为.

9. 考虑系统

$$\begin{aligned} r' &= r - r^2 \\ \theta' &= \sin\theta + a. \end{aligned}$$

(a) 当 a 取哪些值时, 系统会发生分岔?

(b) 描述在分岔值附近 (分岔时、分岔前和分岔后) 解的局部行为.

(c) 简略地作出系统所有可能的、不同的相图.

(d) 讨论在所有分岔中出现的大范围变化.

10. 设 $\boldsymbol{X}' = F(\boldsymbol{X})$ 是 $\mathbb{R}^n$ 中的一个非线性系统. 假设 $F(0) = 0$, 且 DF_0 有 n 个具有负实部的不同特征值. 描述该系统和它的线性化系统之间的一个共轭的构造.

11. 考虑系统 $\boldsymbol{X}' = F(\boldsymbol{X})$, 其中 $\boldsymbol{X} \in \mathbb{R}^n$. 假设系统在 $\boldsymbol{X}_0$ 处有一个平衡点. 证明存在一个坐标变换将 $\boldsymbol{X}_0$ 移到原点, 且将系统化成

$$\boldsymbol{X}' = \boldsymbol{AX} + G(X),$$

其中 $\boldsymbol{A}$ 为一 $n \times n$ 矩阵, 正好就是 DF_{X_0} 的标准形, 而且 $G(\boldsymbol{X})$ 满足

$$\lim_{|\boldsymbol{X}| \to 0} \frac{|G(\boldsymbol{X})|}{|\boldsymbol{X}|} = 0.$$

12. 在定义平衡点为渐近稳定时, 我们要求平衡点也是稳定的. 这个要求不是多余的. 给出一个相图的例子 (草图就行), 其中有一个平衡点, 它附近的所有解曲线 (最终) 都趋于它, 但这个平衡点不是稳定的.

第 9 章　大范围的非线性技巧

在本章, 我们将展示许多分析非线性微分方程系统的行为的定性技巧. 但我们必须事先提醒读者, 这些技巧没有一个对所有非线性系统都适用; 大多数的技巧都只适用于一些特殊情形, 尽管如此, 我们在随后的几章中将看到, 这些特定情形在微分方程的许多重要应用中都会出现.

9.1　零　点　集

零点集 (nullcline)是分析非线性微分方程系统 (特别是平面系统) 的最有用的工具之一. 对于如下形式的系统

$$\begin{aligned} x_1' &= f_1(x_1,\cdots,x_n) \\ &\vdots \\ x_n' &= f_n(x_1,\cdots,x_n), \end{aligned}$$

它的 x_j 零点集是指 x_j' 为零的点构成的集合, 即, 它的 x_j 零点集就是由方程 $f_j(x_1,\cdots,x_n)=0$ 所确定的点集.

x_j 零点集常常将 $\mathbb{R}^n$ 分割成一些区域, 在每个区域上, 该向量场的 x_j 分量要么指向正方向要么指向负方向. 如果确定了所有的零点集, 那么我们就可以将 $\mathbb{R}^n$ 分解成一些区域, 使得在每个区域上, 向量场指向一个“确定的方向”.

在平面系统

$$\begin{aligned} x' &= f(x,y) \\ y' &= g(x,y) \end{aligned}$$

的情形, 上述想法是最容易理解的. 在 x 零点集上, 我们有 $x'=0$, 于是向量场或者垂直向上, 或者垂直向下, 而且只有在零点集上才会有这种情况. 于是, x 零点集将 $\mathbb{R}^2$ 分割成一些区域, 在每个区域中, 向量场或者指向左, 或者指向右. 类似地, 在 y 零点集上, 向量场为水平的, 于是, y 零点集将 $\mathbb{R}^2$ 分割成一些区域, 在每个区域中, 向量场或者指向上, 或者指向下. x 零点集与 y 零点集的交点就是平衡点. 在这两个零点集围成的任何区域中, 向量场既不是竖直的, 也不是水平的, 因而它的指向必定是下面四个方向之一: 东北、西北、东南、西南. 我们称这样的区域为**基本区域**. 常常是这样, 简单勾画出基本区域就可以让我们完全了解相图, 至少从定性的观点看是如此.

例 对于系统

$$\begin{aligned} x' &= y - x^2 \\ y' &= x - 2, \end{aligned}$$

x 零点集为抛物线 $y = x^2$, y 零点集为竖直线 $x = 2$. 这两个零点集相交于 $(2, 4)$, 它就是系统的唯一的平衡点. 这两个零点集将 $\mathbb{R}^2$ 分成从 A 到 D 的四个基本区域, 如图 9.1a 所示. 先在这些区域中各取一点, 然后确定向量场在这一点处的方向, 则我们就可以知道这个基本区域中所有点处向量场的方向. 例如, 点 $(0, 1)$ 位于区域 A, 向量场在该点为 $(1, -2)$, 它指向东南, 于是, 向量场在这个区域中的所有点处都指向东南. 当然, 在这个区域中, 向量场可以是几乎水平的或几乎竖直的, 我们说东南是指向量场的角度 θ 位于扇形 $-\pi/2 < \theta < 0$ 中. 我们可以继续用这种方法得到四个区域中向量场的方向, 如图 9.1b 所示. 这同时也确定了两个零点集上向量场的水平指向和竖直指向.

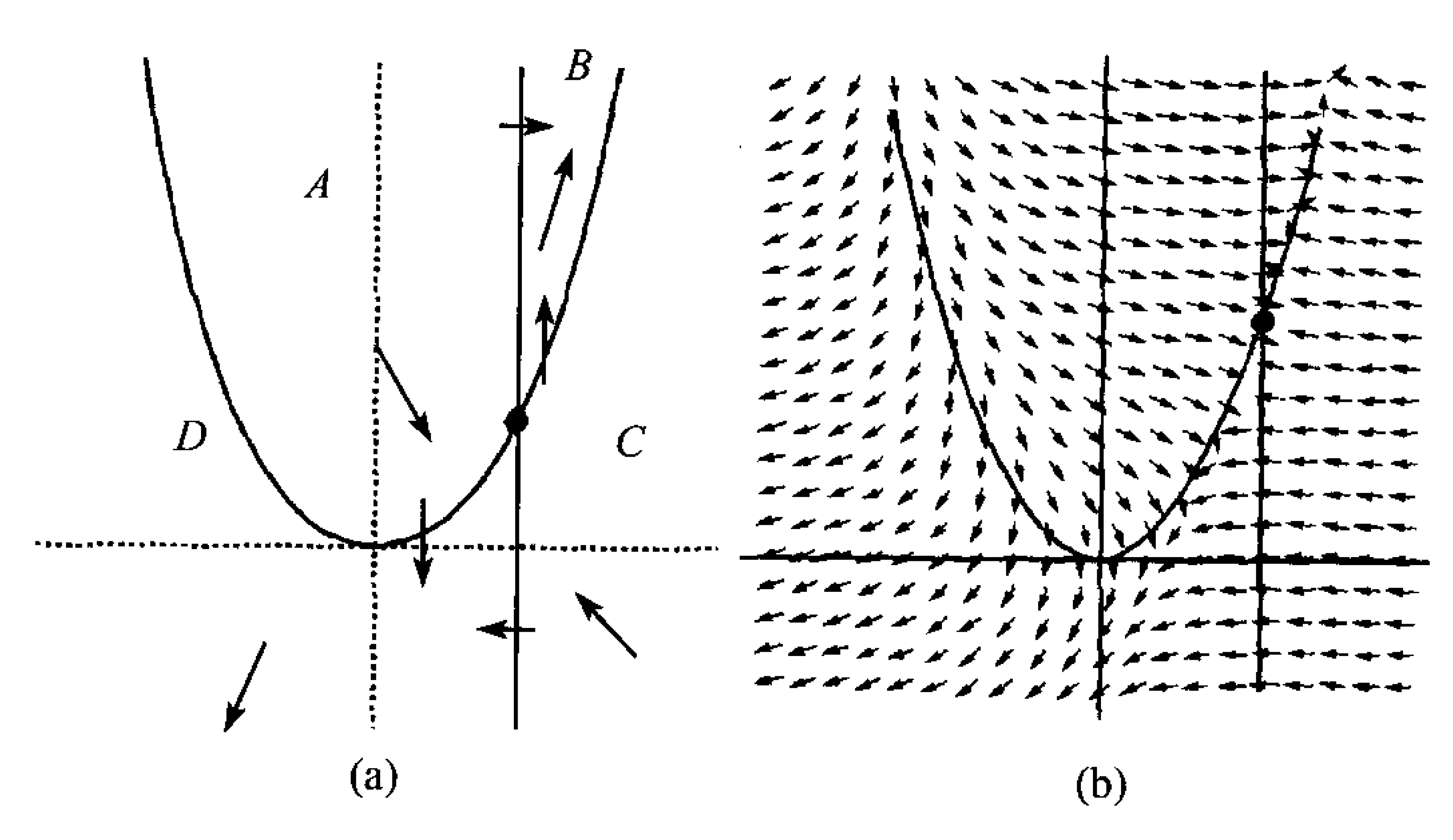


图 9.1 (a) 零点集, (b) 方向场

似乎只从方向场就可以发现, 系统的唯一平衡点为鞍点. 事实上, 的确如此. 因为在平衡点 $(2, 4)$ 处的线性化系统为

$$\boldsymbol{X}' = \begin{pmatrix} -4 & 1 \\ 1 & 0 \end{pmatrix} \boldsymbol{X},$$

它的特征值为 $-2 \pm \sqrt{5}$, 正好一正一负.

更重要的是, 我们可以在全平面上处处给出解的大致行为. 例如, 注意到在标号为 B 的基本区域的所有边界点上, 向量场都是指向该区域的内部, 而在 B 的内部所有点处, 向量场都指向东北. 因而, 区域 B 中的所有解在将来都始终停留在 B 中, 且沿东北方向趋于 ∞(见图 9.2). 类似地, 基本区域 D 中的所有解都停留在该区域中且沿西南方向趋于 ∞. 在基本区域 A 和 C 中出发的解则有两种可能：要

么穿过一个零点集而进入区域 B 和 D(于是我们就知道它的最终行为), 要么趋于平衡点. 然而, 在每个区域中只有一条由这样的解构成的曲线, 即 $(2,4)$ 处的稳定曲线. 这样, 至少从定性的观点, 我们完全了解了该系统的相图 (见图 9.3). ■

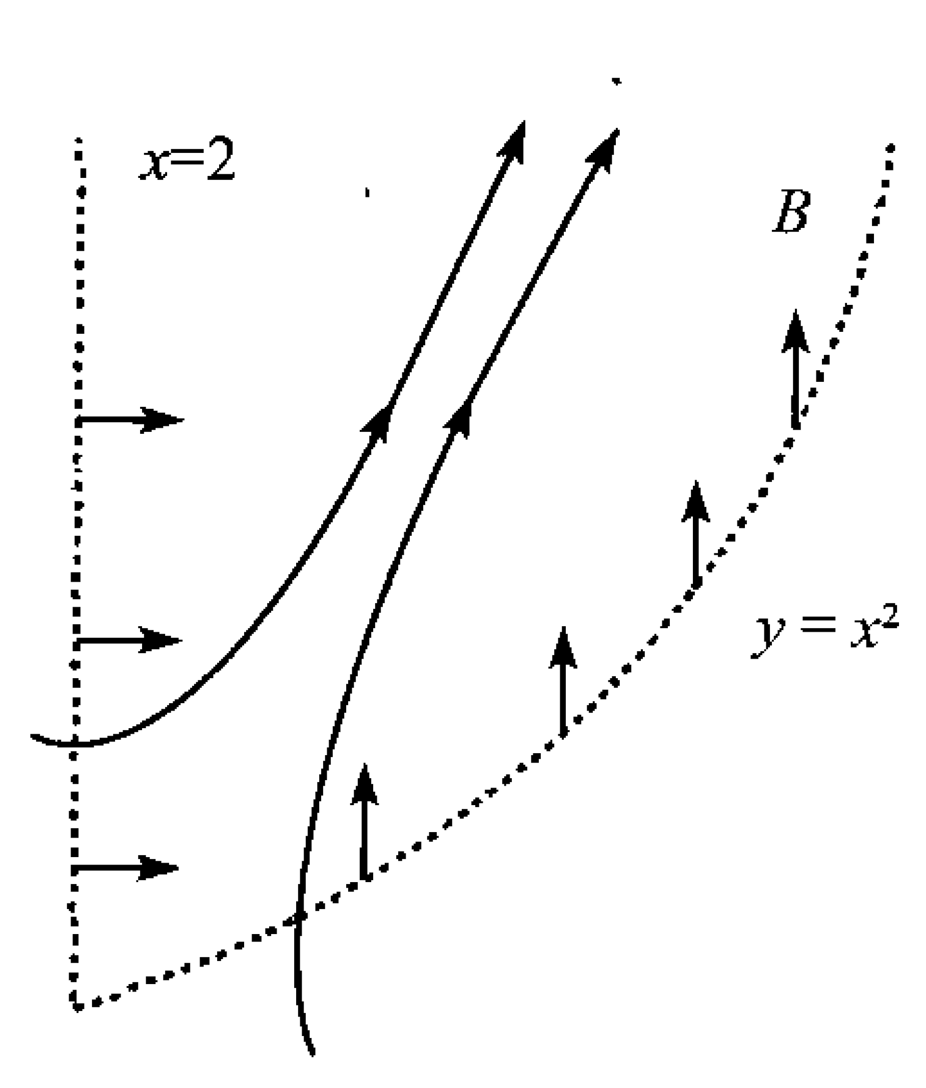


图 9.2 解进入基本区域 B 后趋于 ∞

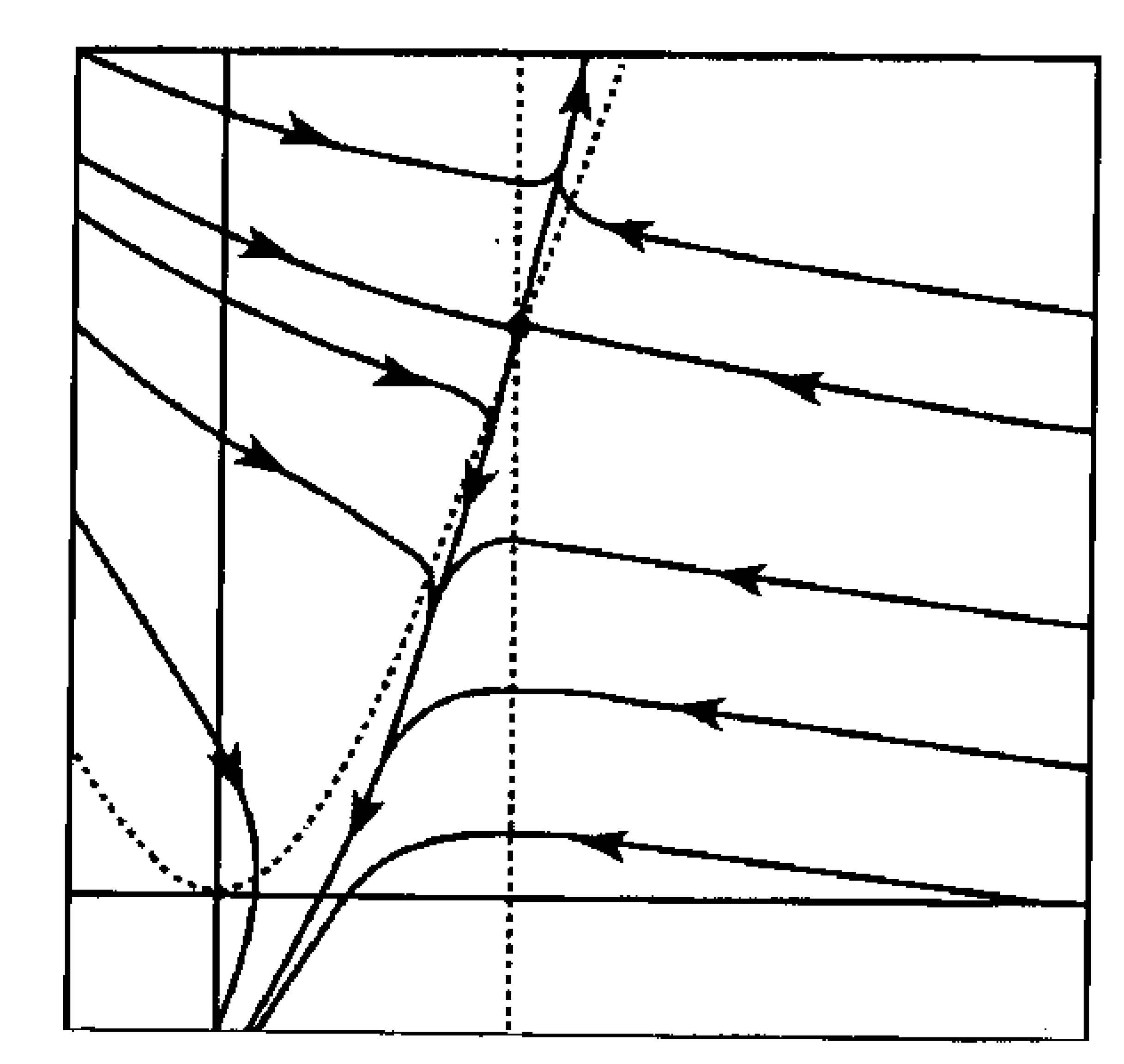


图 9.3 $x' = y - x^2, y' = x - 2$ 的零点集和相图

例 (**异宿分岔**) 下面考虑依赖于参数 a 的系统:

$$x' = x^2 - 1$$

$$y' = -xy + a(x^2 - 1).$$

x 零点集由 $x = \pm 1$ 给出, 而 y 零点集则由 $xy = a(x^2 - 1)$ 给出. 平衡点为 $(\pm 1, 0)$. 由于在 $x = \pm 1$ 上, $x' = 0$, 向量场实际上与这个零点集是相切的. 进一步, 在 $x = 1$ 上有 $y' = -y$, 在 $x = -1$ 上有 $y' = y$. 从而, 沿着竖直线 $x = 1$, 解趋于 $(1, 0)$; 而沿着竖直线 $x = -1$, 解远离 $(-1, 0)$. 对于 a 的所有取值都是如此.

现在, 让我们来看 $a = 0$ 的情形. 此时, 系统简化为

$$x' = x^2 - 1$$

$$y' = -xy,$$

可见沿坐标轴有 $y' = 0$. 特别地, 向量场与 x 轴相切, 且在这条线上为 $x' = x^2 - 1$. 从而, 当 $|x| > 1$ 时, $x' > 0$, 而当 $|x| < 1$ 时, $x' < 0$. 于是, 在每个平衡点处, 既有一个趋于它的直线解也有一个远离它的直线解. 于是, 每个平衡点都似乎是一个鞍点. 用线性化容易验证, 事情的确如此.

y 零点集的另一部分为 $x = 0$, 但此时向量场与零点集不相切. 通过计算由零点集所确定的每个基本区域中的向量场的方向, 可以得到图 9.4, 从中我们可以立刻推断出所有解的定性行为.

注意, 当 $a = 0$ 时, 过 $(1, 0)$ 的不稳定曲线的一支与过 $(-1, 0)$ 的稳定曲线的一

支完全重合. 这条曲线上的所有的解简单地从一个鞍点跑到另一个鞍点. 这样的解称为**异宿解**或**鞍点连线**. 对于平面系统, 稳定和不稳定曲线如此相交形成异宿"连接"的情形是很少见的. 然而, 一旦有了鞍点连线, 人们就可期望出现分岔.

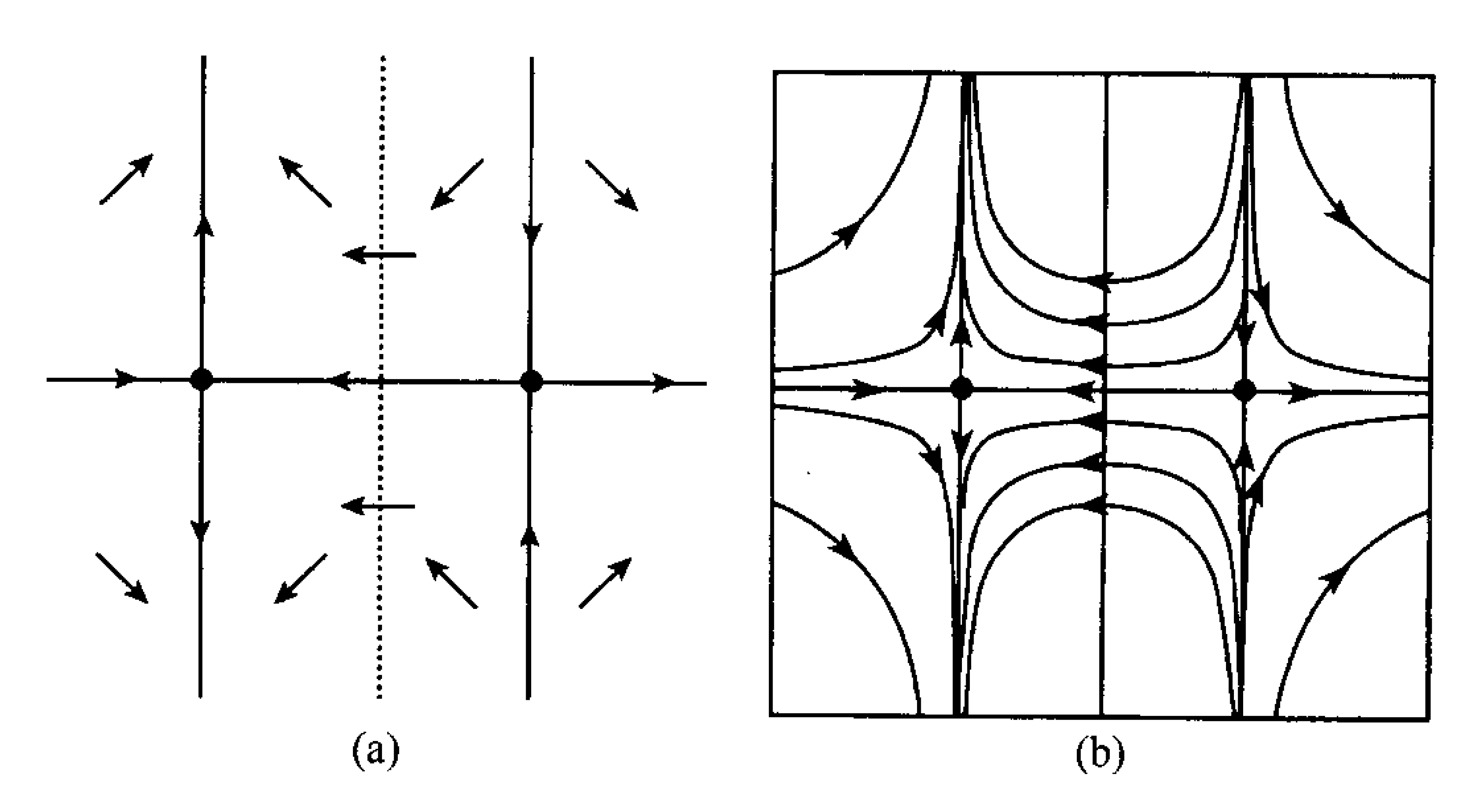


图 9.4　$x' = x^2 - 1, y' = -xy$ 的 (a) 零点集和 (b) 相图

现在考虑 $a \neq 0$ 的情形. x 零点集与 $a = 0$ 时相同, 依然在 $x = \pm 1$ 上. y 零点集却发生了显著的变化, 如图 9.5 所示, 它由 $y = a(x^2 - 1)/x$ 所确定.

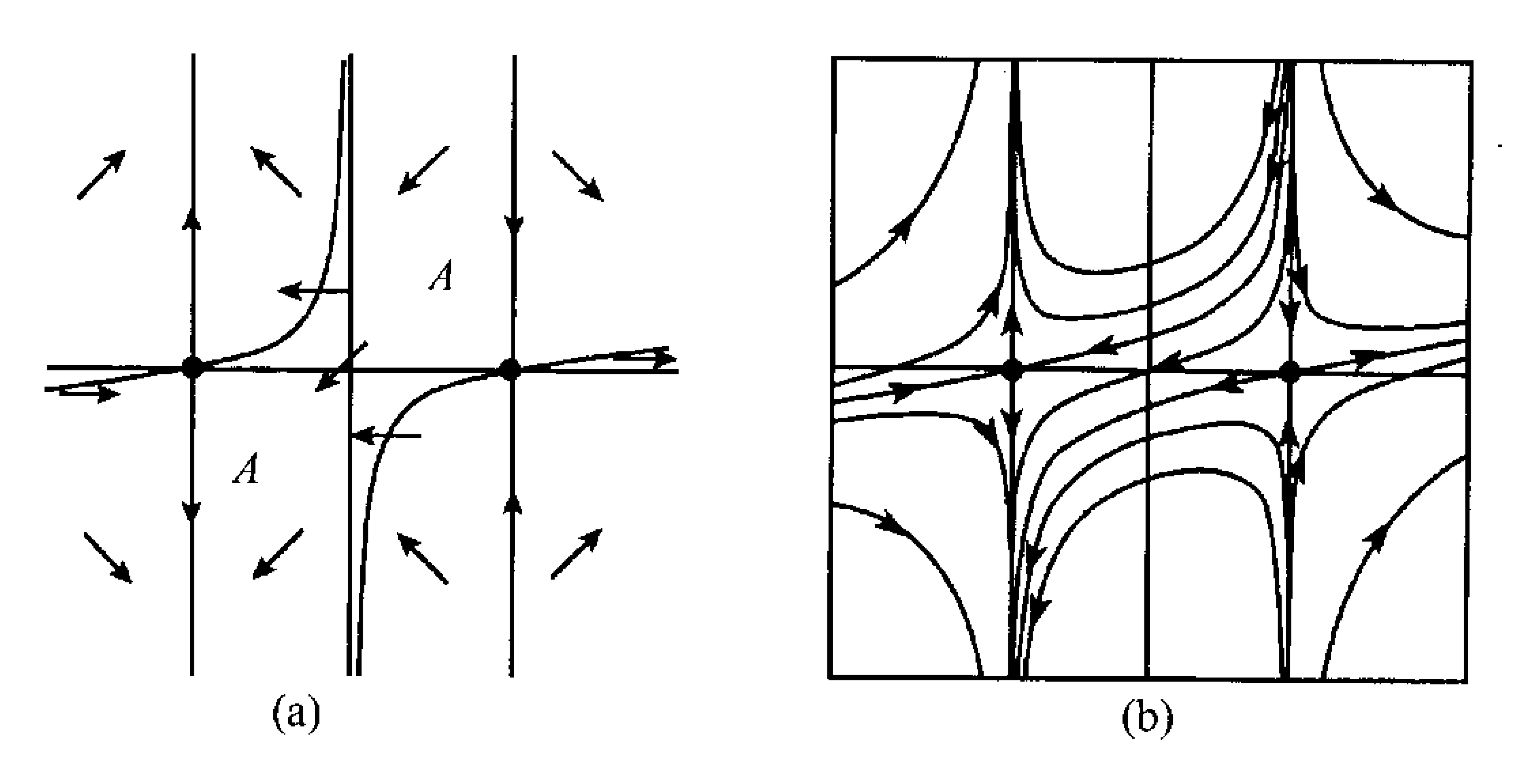


图 9.5　异宿分岔后的 $(a > 0)$ (a) 零点集和 (b) 相平面

当 $a > 0$ 时, 考虑图 9.5 中的基本区域 A, 其中的向量场指向西南方向. 特别地, 在 x 轴上的 $x = -1$ 和 $x = 1$ 之间, 向量场是指向这个方向. 这就破坏了异宿连接. $(-1, 0)$ 的稳定曲线的右边部分必定是从上半平面的 $y = \infty$ 处过来, 而 $(1, 0)$ 的不稳定曲线的左边部分将下降到下半平面的 $y = -\infty$ 处. 这就在两条线 $x = \pm 1$ 之间开辟了一条"大路"使得某些解可由此从 ∞ 跑向 $-\infty$. 而当 $a = 0$ 时, 所有的解要么局限在上半平面, 要么局限在下半平面, 在 $a = 0$ 处发生的**异宿分岔**则打开了一扇门, 让某些解可以从此穿过.

当 $a < 0$ 时将出现类似的情形 (见本章的习题 2).　■

9.2 平衡点的稳定性

当平衡点为双曲时, 可以直接确定它的稳定性, 否则的话, 确定稳定性就很成问题了. 在本节, 我们详细地阐述另一种方法, 用来证明平衡点为渐近稳定的. 这个方法是属于俄罗斯数学家李雅普诺夫的. 在一个具有标准形式的线性系统中, 径向分量 r 沿解曲线递减, 这个方法推广了该想法. 李雅普诺夫注意到 r 之外的其他一些函数也可以达到这一目的. 也许更重要的是, 李雅普诺夫方法可以让我们对一个渐近稳定的平衡点的**吸引盆**的大小有所了解. 根据定义, 吸引盆是指所有那些初始条件构成的集合, 从这些初始条件出发的解都将趋于平衡点.

令 $\mathcal{O}$ 为 $\mathbb{R}^n$ 中一开集, 它包含系统 $\boldsymbol{X}' = F(\boldsymbol{X})$ 的一个平衡点 $\boldsymbol{X}^*$, $L : \mathcal{O} \to \mathbb{R}$ 是一个定义在 $\mathcal{O}$ 上的可微函数. 考虑函数

$$\dot{L}(\boldsymbol{X}) = DL_{\boldsymbol{X}}(F(\boldsymbol{X})).$$

我们已经看到, 如果 $\phi_t(\boldsymbol{X})$ 为该系统在 $t=0$ 时经过 $\boldsymbol{X}$ 的解, 则由链式法则可得

$$\dot{L}(\boldsymbol{X}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} L \circ \phi_t(\boldsymbol{X}).$$

于是, 如果 $\dot{L}(\boldsymbol{X})$ 为负, 则 L 沿过 $\boldsymbol{X}$ 的解曲线递减.

现在我们就能够陈述李雅普诺夫稳定性定理:

定理 (**李雅普诺夫稳定性**) 设 $\boldsymbol{X}^*$ 为 $\boldsymbol{X}' = F(\boldsymbol{X})$ 的一个平衡点, $L : \mathcal{O} \to \mathbb{R}$ 为一可微函数, 其中 $\mathcal{O}$ 为一包含 $\boldsymbol{X}^*$ 的开集. 进一步假设

(a) $L(\boldsymbol{X}^*) = 0$, 且当 $\boldsymbol{X} \neq \boldsymbol{X}^*$ 时, $L(\boldsymbol{X}) > 0$.

(b) 在 $\mathcal{O} - \boldsymbol{X}^*$ 中, $\dot{L} \leqslant 0$.

则 $\boldsymbol{X}^*$ 是稳定的. 更进一步, 如果 L 还满足

(c) 在 $\mathcal{O} - \boldsymbol{X}^*$ 中, $\dot{L} < 0$,

则 $\boldsymbol{X}^*$ 是渐近稳定的. ■

满足 (a) 和 (b) 的函数 L 称为 $\boldsymbol{X}^*$ 的一个**李雅普诺夫函数**. 如果 (c) 也满足, 我们称 L 为一个**严格**李雅普诺夫函数.

注意, 应用李雅普诺夫定理时可以不用求解微分方程, 我们唯一需要做的是计算 $DL_{\boldsymbol{X}}(F(\boldsymbol{X}))$. 这点非常有用! 另一方面, 不存在现成的求李雅普诺夫函数的方法. 通常在每种情形都有一些特别的技巧, 或者需要反复摸索. 有时, 可以用一些自然的函数来试探. 例如, 我们在第 13 章将看到, 在力学或电学系统情形, 能量常常就是一个李雅普诺夫函数.

例 考虑 $\mathbb{R}^3$ 中微分方程系统

$$\begin{aligned}x' &= (\epsilon x + 2y)(z+1)\\ y' &= (-x+\epsilon y)(z+1)\\ z' &= -z^3,\end{aligned}$$

其中 ϵ 为一个参数. 原点是该系统的唯一平衡点. 系统在 $(0,0,0)$ 处的线性化为

$$\boldsymbol{Y}' = \begin{pmatrix} \epsilon & 2 & 0 \\ -1 & \epsilon & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{Y}.$$

它的特征值为 0 和 $\epsilon \pm \sqrt{2}\mathrm{i}$. 因而, 从线性化我们只能得出当 $\epsilon > 0$ 时, 原点是不稳定的. 当 $\epsilon \leqslant 0$ 时, 我们只能得出原点不是双曲的.

当 $\epsilon \leqslant 0$ 时, 我们来寻找 $(0,0,0)$ 的一个形如 $L(x,y,z) = ax^2 + by^2 + cz^2$ 的李雅普诺夫函数, 其中 $a, b, c > 0$. 对于这样的 L, 我们有

$$\dot{L} = 2(axx' + byy' + czz'),$$

从而

$$\begin{aligned}\dot{L}/2 =& \ ax(\epsilon x + 2y)(z+1) + by(-x+\epsilon y)(z+1) - cz^4\\ =& \ \epsilon(ax^2+by^2)(z+1) + (2a-b)(xy)(z+1) - cz^4.\end{aligned}$$

为了得到稳定性, 我们需要 $\dot{L} \leqslant 0$; 这是可以达到的, 例如, 令 $a=1, b=2, c=1$. 当 $\epsilon = 0$ 时, 我们有 $\dot{L} = -z^4 \leqslant 0$, 从而原点是稳定的. 可以证明在这种情形, 原点不是渐近稳定的 (见习题 4).

当 $\epsilon < 0$ 时, 我们有

$$\dot{L} = \epsilon(x^2+2y^2)(z+1) - z^4,$$

于是在由 $z > -1$ 定义的区域 $\mathcal{O}$ 中, $\dot{L} < 0$(除去原点). 由此可得, 在这种情形, 原点是渐近稳定的. 事实上, 由习题 4 知, 所有从区域 $\mathcal{O}$ 中出发的解都趋于原点. ■

例 (**非线性单摆**) 考虑由一长为 l 的轻杆和一质量为 m 的小球 (看成质点) 相连构成的单摆. 假设杆的另一端与一墙面上的一点相连, 从而单摆的球沿以该点为中心的圆周运动. 质点在 t 时刻的位置可由它离开铅垂位置的角度 $\theta(t)$ 所完全确定, 其中角度 $\theta(t)$ 沿反时针方向测量. 于是质点在 t 时刻的位置为 $(l\sin\theta(t), -l\cos\theta(t))$.

质点的速率就是速度向量的长度, 即 $l\mathrm{d}\theta/\mathrm{d}t$, 其加速度为 $l\mathrm{d}^2\theta/\mathrm{d}t^2$. 我们假设单摆上的作用力只有重力和摩擦力. 重力是常值的作用力, 大小为 mg, 方向向下; 这个力在切于运动圆周的方向上的分量为 $-mg\sin\theta$. 我们认为摩擦力与速度成比

例, 从而这个力由 $-bl\mathrm{d}\theta/\mathrm{d}t$ 给出, 其中 b 为某常数. 当没有摩擦力时 $(b=0)$, 我们得到**理想单摆**. 于是牛顿定律就给出单摆的二阶微分方程

$$ml\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}=-bl\frac{\mathrm{d}\theta}{\mathrm{d}t}-mg\sin\theta.$$

为了简单起见, 我们假设选取单位使得 $m=l=g=1$. 引入 $v=\mathrm{d}\theta/\mathrm{d}t$ 后可将这个方程改写成系统

$$\begin{aligned}\theta'&=v\\ v'&=-bv-\sin\theta.\end{aligned}$$

显然, 该系统有两个平衡点 $(\mathrm{mod}2\pi)$: 下方的静止位置 $\theta=0, v=0$, 以及垂直向上的位置 $\theta=\pi, v=0$. 无论是从数学上看 (检查其线性化), 还是从物理上看, 上方的位置都是一个不稳定的平衡点.

对于下方的平衡点, 其线性化方程为

$$\boldsymbol{Y}'=\begin{pmatrix}0&1\\-1&-b\end{pmatrix}\boldsymbol{Y}.$$

它的特征值要么是纯虚的 (当 $b=0$ 时), 要么具有负实部 (当 $b>0$ 时). 从而当 $b>0$ 时, 下方的平衡点是渐近稳定的. 地球人, 只要是见过真实的单摆, 都知道这点.

为了进一步研究这一个平衡点, 考虑函数 $E(\theta,v)=\dfrac{1}{2}v^2+1-\cos\theta$. 对于有初等力学背景的读者来说, 这就是熟知的**总能量**函数, 我们将在第 13 章对它作深入描述. 我们来计算

$$\dot{E}=vv'+\sin\theta\,\theta'=-bv^2,$$

从而 $\dot{E}\leqslant 0$. 于是 E 是一个李雅普诺夫函数, 而且原点是一个稳定的平衡点. 如果 $b=0$(即没有摩擦), 则 $\dot{E}\equiv 0$. 这说明, E 沿着该系统所有的解都是常值. 于是, 只须简单地作出 E 的水平曲线就可看出解曲线的位置. 该系统的相图如图 9.6 所示. 注意, 我们并不用求解该微分方程就可画出相图; 知道了 E 的水平曲线 (和向量场的方向), 一切也就清楚了. 在本章的后面, 我们将会碰到许多这样 (非常特殊) 的函数, 它们沿解曲线都是常值.

环绕原点的解都具有性质: 对所有的 t, $-\pi<\theta<\pi$. 因而, 这些解对应于单摆关于下方的静止位置来回摆动, 而不会穿过上面的位置 $\theta=\pi$. 连接位于 $(\pm\pi,0)$ 的平衡点的两个特殊解分别对应于单摆在时间向前和向后时趋于上方平衡点. (你不大常见到这种运动!) 除了这些特殊的解外, 我们发现对于其他的解, $\theta(t)$ 或者永远增加或者永远减少; 这分别对应于单摆永远反时针旋转或顺时针旋转. ■

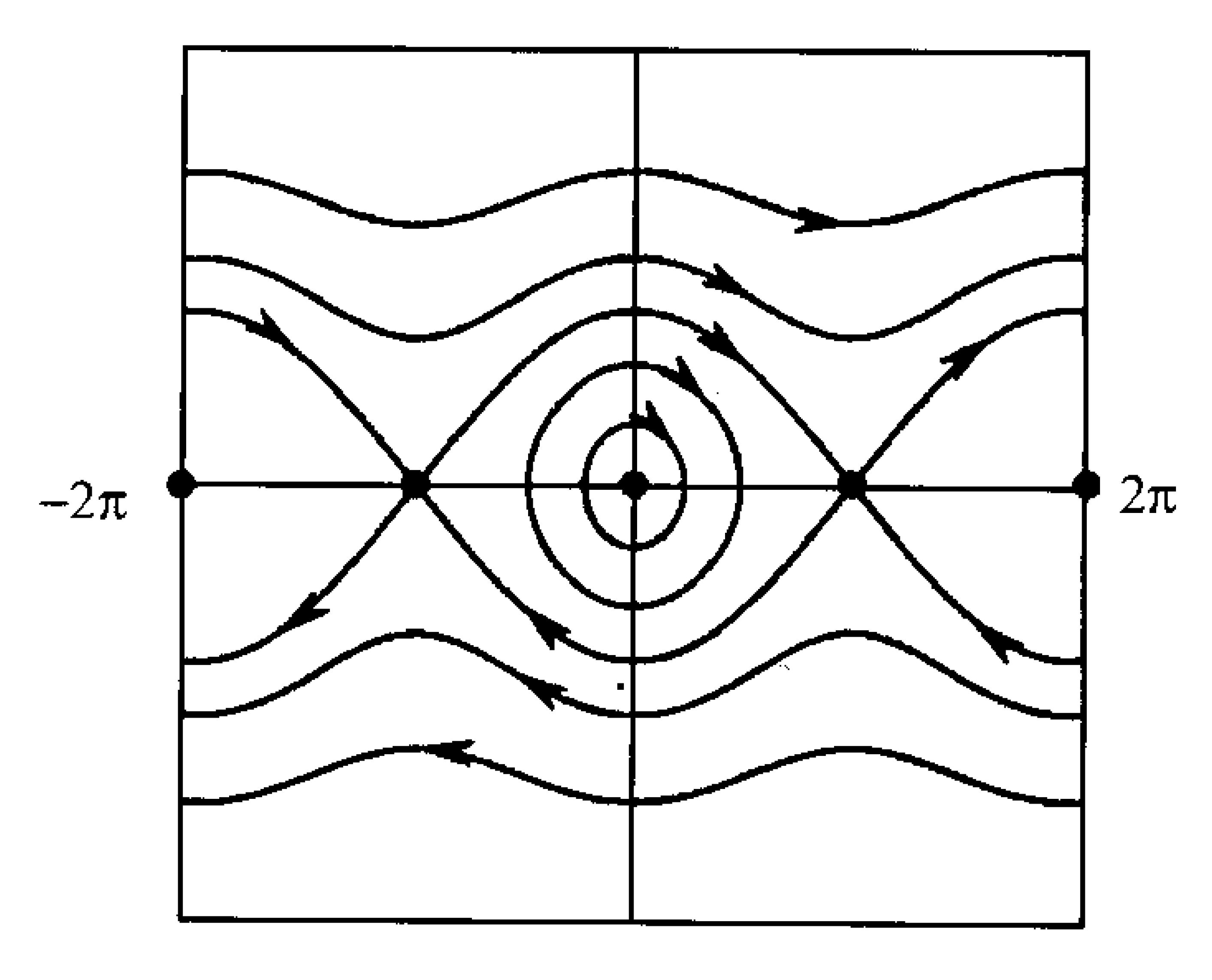


图 9.6　理想单摆的相图

在后面, 我们将会回到单摆在 $b > 0$ 的情形. 但是, 我们先来证明李雅普诺夫定理.

证明　令 δ 是如此的小, 以至于围绕平衡点 $\boldsymbol{X}^*$ 的半径为 δ 的闭球 $B_\delta(\boldsymbol{X}^*)$ 整个位于 $\mathcal{O}$ 中. 令 α 为 L 在 $B_\delta(\boldsymbol{X}^*)$ 的边界上的最小值, 其中 $B_\delta(\boldsymbol{X}^*)$ 的边界就是中心在 $\boldsymbol{X}^*$, 半径为 δ 的球面 $S_\delta(\boldsymbol{X}^*)$. 于是, 由假设, $\alpha > 0$. 令 $\mathcal{U} = \{\boldsymbol{X} \in B_\delta(\boldsymbol{X}^*)|L(\boldsymbol{X}) < \alpha\}$, 注意 $\boldsymbol{X}^*$ 位于 $\mathcal{U}$ 中. 由于 L 沿解曲线不增, 从 $\mathcal{U}$ 中出发的解不会与 $S_\delta(\boldsymbol{X}^*)$ 相交. 于是, 每个从 $\mathcal{U}$ 中出发的解永远都不会离开 $B_\delta(\boldsymbol{X}^*)$. 这就证明了 $\boldsymbol{X}^*$ 是稳定的.

现在假设李雅普诺夫稳定性定理中的条件 (c) 也满足, 从而, 沿 $\mathcal{U} - \boldsymbol{X}^*$ 中的解, L 是严格减少的. 令 $\boldsymbol{X}(t)$ 是一个从 $\mathcal{U} - \boldsymbol{X}^*$ 中出发的解, 且假设当某个序列 $t_n \to \infty$ 时, $\boldsymbol{X}(t_n) \to \boldsymbol{Z}_0 \in B_\delta(\boldsymbol{X}^*)$. 我们断言 $\boldsymbol{Z}_0 = \boldsymbol{X}^*$. 为此, 首先观察到, 由于 $L(\boldsymbol{X}(t))$ 减小, 因而对于所有的 $t \geqslant 0$, $L(\boldsymbol{X}(t)) > L(\boldsymbol{Z}_0)$, 再由 L 的连续性可得 $L(\boldsymbol{X}(t_n)) \to L(\boldsymbol{Z}_0)$. 如果 $\boldsymbol{Z}_0 \neq \boldsymbol{X}^*$, 令 $\boldsymbol{Z}(t)$ 为从 $\boldsymbol{Z}_0$ 出发的解. 对任何 $s > 0$, 我们有 $L(\boldsymbol{Z}(s)) < L(\boldsymbol{Z}_0)$. 于是, 对于任一从充分接近 $\boldsymbol{Z}_0$ 出发的解, 我们都有

$$L(\boldsymbol{Y}(s)) < L(\boldsymbol{Z}_0).$$

取 $\boldsymbol{Y}(0) = \boldsymbol{X}(t_n)$, 其中 n 充分大, 则我们就得到一个矛盾

$$L(\boldsymbol{X}(t_n + s)) < L(\boldsymbol{Z}_0).$$

因而, $\boldsymbol{Z}_0 = \boldsymbol{X}^*$. 这就证明了 $\boldsymbol{X}^*$ 是集合 $\{\boldsymbol{X}(t)|t > 0\}$ 唯一可能的极限点, 从而完成了李雅普诺夫定理的证明. ■

图 9.7 使得该定理直观上看起来显然. 条件 $\dot{L} < 0$ 意味着, 当一个解穿过“水平曲面” $L^{-1}(c)$ 后, 它就跑到了集合 $L \leqslant c$ 的内部, 并且再也不能跑出来. 遗憾的是, 图中所示意的时常很难验证; 为什么集合 $L^{-1}(c)$ 要缩小到 $\boldsymbol{X}^*$? 当然, 在许多情形, 图 9.7 事实上是正确的, 例如, 当 L 为 $ax^2 + by^2 (a, b > 0)$ 这样的二次函数时.

但是, 如果水平曲面看起来像图 9.8 呢? 很难想象这样的 L 能够满足李雅普诺夫函数的所有要求; 但是与其去排除这种可能性, 还不如像上面那样给出分析证明来得简单.

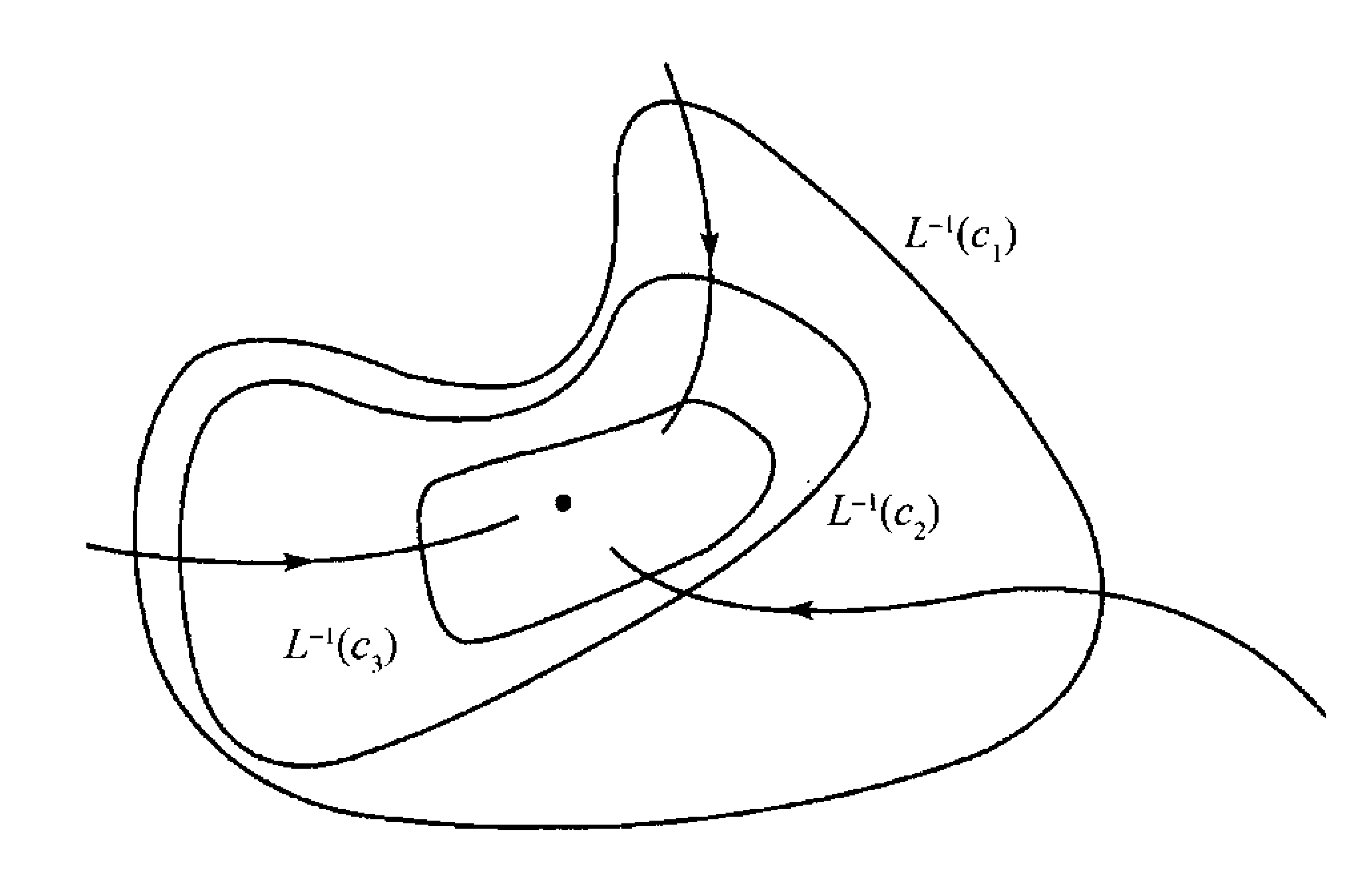


图 9.7 解递减地穿过一严格李雅普诺夫函数的水平集 $L^{-1}(c_j)$

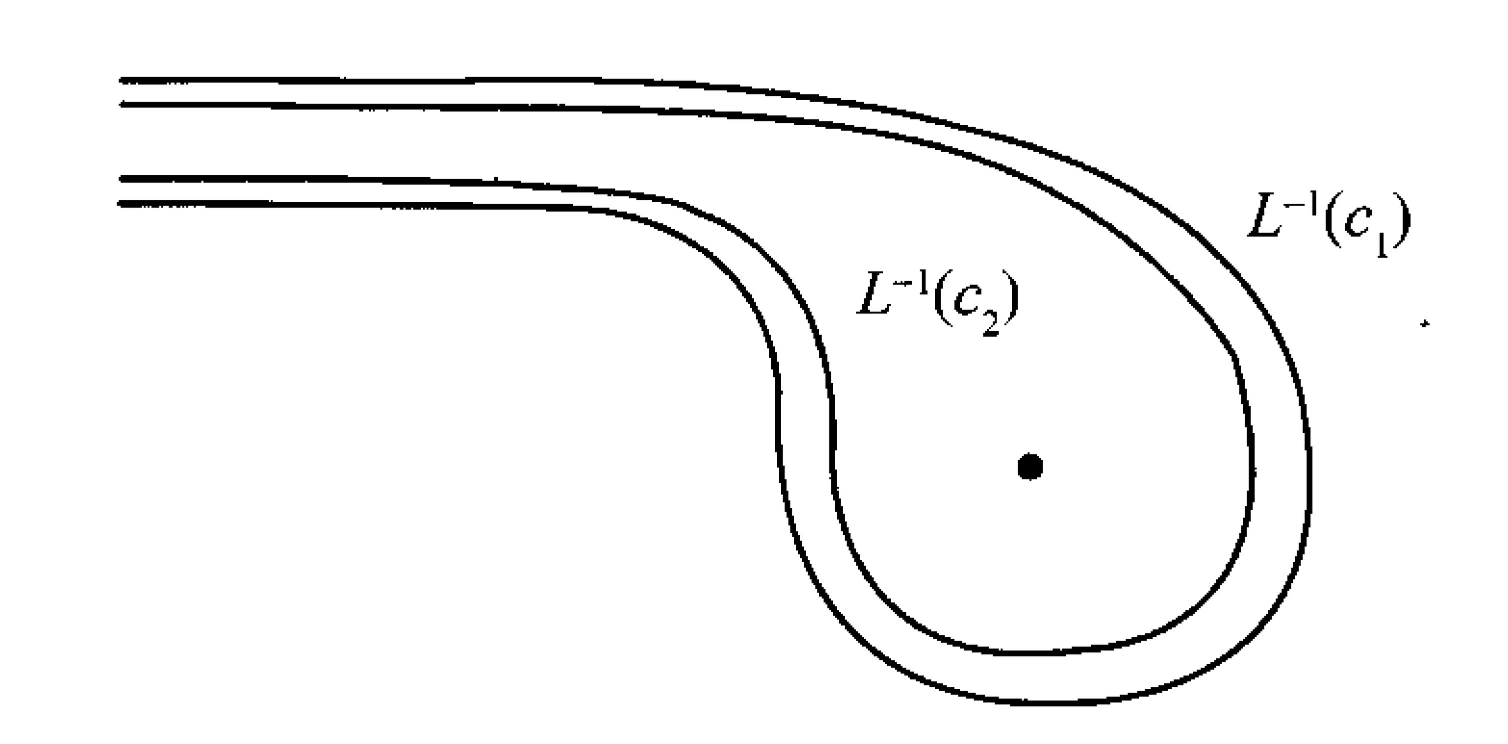


图 9.8 李雅普诺夫函数的水平集有可能看起来像这样

例 现在考虑系统

$$\begin{aligned} x' &= -x^3 \\ y' &= -y(x^2+z^2+1) \\ z' &= -\sin z. \end{aligned}$$

原点又是一个平衡点, 然而, 并不是只有这一个平衡点, 事实上, 对于每个 $n \in \mathbb{Z}$, $(0,0,n\pi)$ 也是一个平衡点. 于是, 原点不可能是大范围渐近稳定的. 而且, 对于每个 $n \in \mathbb{Z}$, 平面 $z = n\pi$ 都是**不变的**, 也就是说, 从这个平面上出发的解将永远逗留在该平面上. 这是因为, 当 $z = n\pi$ 时, 有 $z' = 0$. 特别地, 任何从区域 $|z| < \pi$ 中出发的解都始终逗留在这个区域中.

在原点线性化就得到系统

$$\boldsymbol{Y}' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \boldsymbol{Y},$$

它对于判定平衡点的稳定性毫无用处.

不过, 考虑函数

$$L(x,y,z)=x^2+y^2+z^2.$$

显然, $L>0$(除了在原点例外), 我们来计算

$$\dot{L}=-2x^4-2y^2(x^2+z^2+1)-2z\sin z.$$

于是, 在集合 $|z|<\pi$ 中, 当 $z\neq 0$ 时, $z\sin z>0$, 因而在这个集合的所有点处都有 $\dot{L}<0$ (原点除外). 从而, 原点是渐近稳定的.

进一步, 我们能够推出, 原点的吸引盆就是整个区域 $|z|<\pi$. 从李雅普诺夫稳定性定理的证明中可以立刻得出, 所有从半径 $r<\pi$ 的球面内部出发的解都必定趋于原点. 在半径为 π 的球面外且位于两平面 $z=\pm\pi$ 之间, 函数 L 仍然是严格递减的. 由于解始终陷在这两个平面之间, 因而它们也必定趋于原点. ■

这个例子表明, 李雅普诺夫函数不仅可以用来判定平衡点的稳定性, 而且可以用来估计渐近稳定的平衡点的吸引盆的大小, 正如上例所示. 即使在李雅普诺夫函数不是严格的情形, 我们仍有下面的定理, 它同样给出了渐近稳定性的一个判据, 而且给出了吸引盆的大小. 为了陈述它, 我们需要几个定义.

回忆一下, 一个集合 $\mathcal{P}$ 称为**不变**的, 如果对每一 $\boldsymbol{X}\in\mathcal{P}$, 对所有的 $t\in\mathbb{R}$, $\phi_t(\boldsymbol{X})$ 都有定义且都在 $\mathcal{P}$ 中. 例如, 上例中的区域 $|z|<\pi$ 就是一个不变集. 集合 $\mathcal{P}$ 称为**正向不变**的, 如果对每一 $\boldsymbol{X}\in\mathcal{P}$, 对所有的 $t\geqslant 0$, $\phi_t(\boldsymbol{X})$ 都有定义且都在 $\mathcal{P}$ 中. 在上例的区域 $|z|<\pi$ 中, 一个中心在原点的球面内部就是一个正向不变集, 但不是不变集. 最后, 系统的一个**完整解**是指一个形如 $\{\phi_t(\boldsymbol{X})|t\in\mathbb{R}\}$ 的集合.

定理 (**Lasalle 不变性原理**) 设 $\boldsymbol{X}^*$ 为 $\boldsymbol{X}'=F(\boldsymbol{X})$ 的一个平衡点, $L:\mathcal{U}\to\mathbb{R}$ 为 $\boldsymbol{X}^*$ 的一个李雅普诺夫函数, 其中 $\mathcal{U}$ 是一个包含 $\boldsymbol{X}^*$ 的开集. 令 $\mathcal{P}\subset\mathcal{U}$ 为 $\boldsymbol{X}^*$ 的一个闭有界邻域. 假设 $\mathcal{P}$ 是正向不变的, 而且在 $\mathcal{P}-\boldsymbol{X}^*$ 中没有完整解使得 L 在上面为常值. 则 $\boldsymbol{X}^*$ 是渐近稳定的, 而且 $\mathcal{P}$ 包含在 $\boldsymbol{X}^*$ 的吸引盆中. ■

在证明该定理之前, 我们先将它应用到前面讨论的有阻尼的单摆的平衡点 $\boldsymbol{X}^*=(0,0)$. 回忆一下, 那时的一个李雅普诺夫函数为 $E(\theta,v)=\dfrac{1}{2}v^2+1-\cos\theta$, 且 $\dot{E}=-bv^2$. 由于在 $v=0$ 时, $\dot{E}=0$, 这个李雅普诺夫函数不是严格的.

为了估计 $(0,0)$ 的吸引盆, 固定常数 c, $0<c<2$, 然后定义

$$\mathcal{P}_c=\{(\theta,v)|E(\theta,v)\leqslant c,\ \text{且}\ |\theta|<\pi\}.$$

显然, $(0,0)\in\mathcal{P}_c$. 我们来证明 $\mathcal{P}_c$ 位于 $(0,0)$ 的吸引盆中.

首先注意到 $\mathcal{P}_c$ 是正向不变的. 为了看出这点, 假设 $(\theta(t),v(t))$ 是一个满足 $(\theta(0),v(0))\in\mathcal{P}_c$ 的解. 我们断言, 对于所有的 $t\geqslant 0$, 都有 $(\theta(t),v(t))\in\mathcal{P}_c$. 由于

$\dot{E} \leqslant 0$, 显然有 $E(\theta(t), v(t)) \leqslant c$. 如果 $|\theta(t)| \geqslant \pi$, 则一定存在一个最小的 t_0 使得 $\theta(t_0) = \pm\pi$. 于是,

$$E(\theta(t_0), v(t_0)) = E(\pm\pi, v(t_0)) = \frac{1}{2}v(t_0)^2 + 2 \geqslant 2.$$

然而,

$$E(\theta(t_0), v(t_0)) \leqslant c < 2.$$

这一矛盾就证明了 $|\theta(t_0)| < \pi$, 从而 $\mathcal{P}_c$ 是正向不变的.

现在我们来证明在 $\mathcal{P}_c$ 中不存在完整解使得 E 在上面为常数 (平衡点除外). 假设有这样的一个解, 则沿这个解, $\dot{E} \equiv 0$, 从而 $v \equiv 0$. 于是 $\theta' = 0$, 这意味着 θ 在这个解上为常数. 同时, 在这个解上我们还有 $v' = -\sin\theta = 0$. 由 $|\theta| < \pi$, 可得 $\theta \equiv 0$. 从而, $\mathcal{P}_c$ 中 E 在上面为常数的完整解只能是平衡点 $(0, 0)$.

最后, $\mathcal{P}_c$ 还是一个闭集. 因为, 如果 (θ_0, v_0) 是 $\mathcal{P}_c$ 的一个极限点, 则 $|\theta_0| \leqslant \pi$, 且由 E 的连续性有 $E(\theta_0, v_0) \leqslant c$. 但是, 上面已经证明 $|\theta_0| = \pi$ 蕴涵 $E(\theta_0, v_0) > c$. 于是, $|\theta_0| < \pi$, 故 $(\theta_0, v_0) \in \mathcal{P}_c$, 从而 $\mathcal{P}_c$ 是闭的.

根据上面的定理我们就有, 对于每个 $c < 2$, $\mathcal{P}_c$ 都属于 $(0, 0)$ 的吸引盆; 从而集合

$$\mathcal{P} = \bigcup\{\mathcal{P}_c | 0 < c < 2\}$$

也包含在这个吸引盆中. 注意, 我们可将 $\mathcal{P}$ 写成

$$\mathcal{P} = \{(\theta, v) | E(\theta, v) < 2 \text{ 且 } |\theta| < \pi\}.$$

图 9.9 给出了有阻尼的单摆的相图. 其中的曲线 γ_c 就是水平集 $E(\theta, v) = c$. 注意, 解恰好穿过这些曲线一次并最终趋于原点.

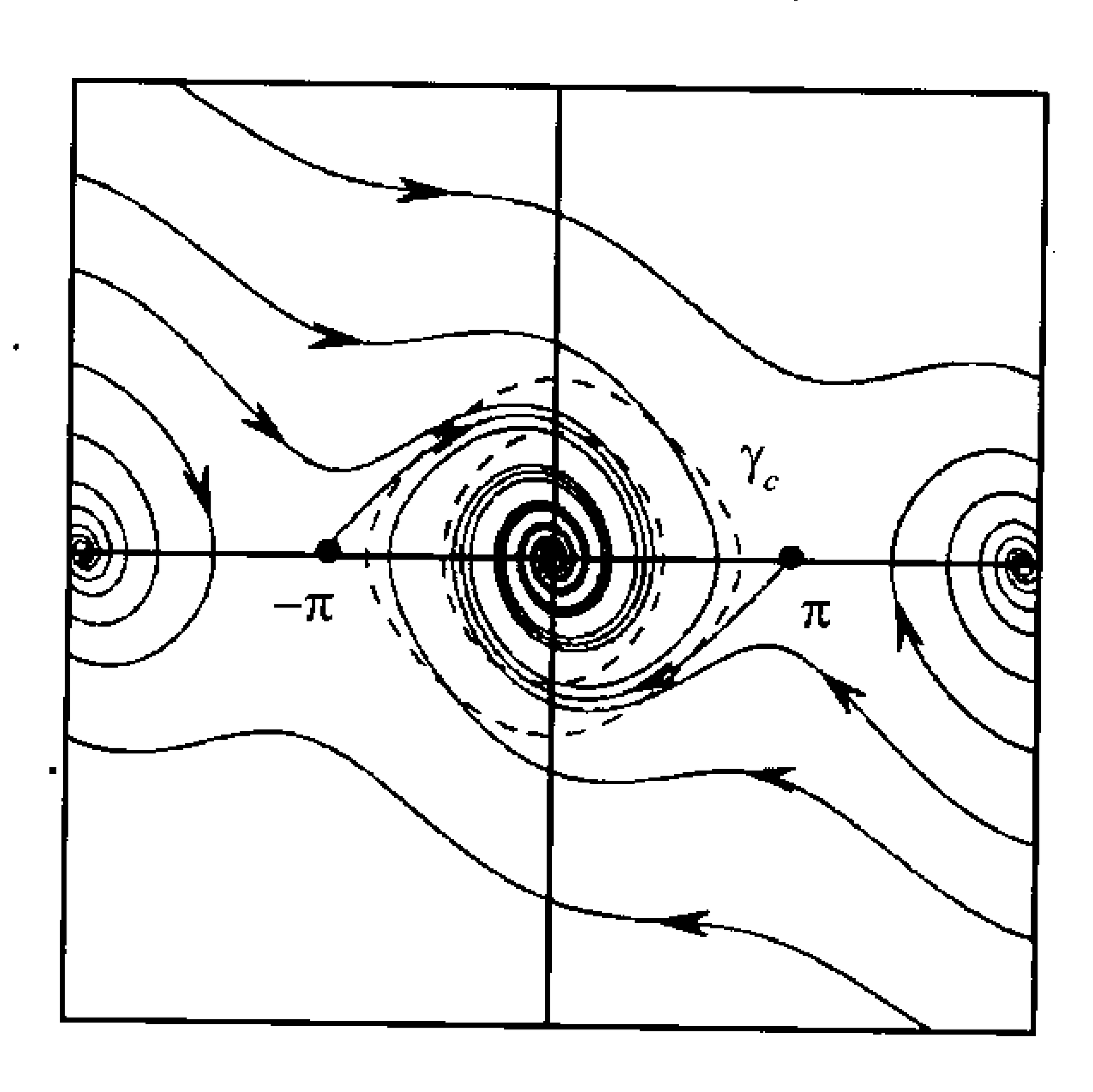


图 9.9 曲线 γ_c 围成区域 $\mathcal{P}_c$

这个结果从物理上看是很自然的. 因为如果 $\theta \neq \pm\pi$, 则 $E(\theta, 0) < 2$, 从而过 $(\theta, 0)$ 的解趋于 $(0,0)$. 即, 如果单摆是从离开静止位置任一角度 θ 的地方 (竖直位置除外) 开始, 则它会越摆越低并最终停在其稳定的平衡位置.

$(0,0)$ 的吸引盆中还包含 $\mathcal{P}$ 之外的其他初始位置. 例如, 考虑过 $(-\pi, u)$ 的解, 其中 u 很小, 但不是零. 则 $(-\pi, u) \notin \mathcal{P}$, 但过这点的解很快就会运动到 $\mathcal{P}$ 中, 从而最终趋于 $(0,0)$. 这说明 $(-\pi, u)$ 也位于 $(0,0)$ 的吸引盆中. 这点可从图 9.9 中看出: 从平衡点 $(-\pi, 0)$ 上方一点点或平衡点 $(\pi, 0)$ 下方一点点的地方出发的解很快就会穿过 γ_c 进入 $\mathcal{P}_c$. 在习题 5 中可以看到进一步的这种例子.

现在我们来证明上述定理.

证明 假设存在一个解 $\boldsymbol{X}(t)$ 在 $0 \leqslant t < \infty$ 时位于正向不变集 $\mathcal{P}$ 中, 但是在 $t \to \infty$ 时, $\boldsymbol{X}(t)$ 不趋于 $\boldsymbol{X}^*$. 由于 $\mathcal{P}$ 是有界闭的, 故一定存在 $\mathcal{P}$ 中的一点 $\boldsymbol{Z} \neq \boldsymbol{X}^*$ 以及一个序列 $t_n \to \infty$ 使得

$$\lim_{n \to \infty} \boldsymbol{X}(t_n) = \boldsymbol{Z}.$$

我们不妨假设 t_n 是一个单调增加序列.

我们断言过 $\boldsymbol{Z}$ 的完整解位于 $\mathcal{P}$ 中, 即, $\phi_t(\boldsymbol{Z})$ 对所有的 $t \in \mathbb{R}$, 而不仅仅是对 $t \geqslant 0$, 都有定义且位于 $\mathcal{P}$ 中. 这可以通过如下的讨论看出. 首先, 由于 $\mathcal{P}$ 是正向不变的, $\phi_t(\boldsymbol{Z})$ 当然对所有的 $t \geqslant 0$ 有定义. 另一方面, 对于区间 $[-t_n, 0]$ 中的所有 t, $\phi_t(\boldsymbol{X}(t_n))$ 都有定义且位于 $\mathcal{P}$ 中. 由于 $\{t_n\}$ 为单调增加序列, 对所有的 $k \geqslant 0$, 以及所有的 $t \in [-t_n, 0]$, $\phi_t(\boldsymbol{X}(t_{n+k}))$ 都有定义且位于 $\mathcal{P}$ 中. 由于当 $k \to \infty$ 时, 点列 $\boldsymbol{X}(t_{n+k}) \to \boldsymbol{Z}$, 根据解对初始条件的连续依赖性, 我们就有, 对于所有的 $t \in [-t_n, 0]$, $\phi_t(\boldsymbol{Z})$ 也有定义且位于 $\mathcal{P}$ 中. 由于这对所有的 t_n 都成立, 可见过 $\boldsymbol{Z}$ 的解是一个位于 $\mathcal{P}$ 中的完整解.

最后, 我们说明在过 $\boldsymbol{Z}$ 的完整解上, L 为常数. 记 $L(\boldsymbol{Z}) = \alpha$, 则我们有 $L(\boldsymbol{X}(t_n)) \geqslant \alpha$ 且

$$\lim_{n \to \infty} L(\boldsymbol{X}(t_n)) = \alpha.$$

更一般地, 只要 $\{s_n\}$ 是一个在 $n \to \infty$ 时趋于 ∞ 的时间序列, 我们就有 $L(\boldsymbol{X}(s_n)) \to \alpha$. 这可以由 L 沿解是非增的事实得出. 现在, 序列 $\boldsymbol{X}(t_n + s)$ 收敛到 $\phi_s(\boldsymbol{Z})$, 从而 $L(\phi_s(\boldsymbol{Z})) = \alpha$. 这与我们假设的 $\mathcal{P}$ 中没有使得 L 取常值的完整解矛盾, 这一矛盾就证明了定理. ■

在这个证明中, 我们遇到了某些点, 它们是过 $\boldsymbol{X}$ 的解上的点列的极限. 一个给定解 $\boldsymbol{X}(t)$ 的所有极限点构成的集合称为这个解的 ω **极限点集**, 或 ω **极限集**. 类似地, 我们可以定义, 一个解 $\boldsymbol{X}(t)$ 的 α **极限点集**, 或 α **极限集**为所有这样的点 $\boldsymbol{Z}$ 构成的集合: 对于某个序列 $t_n \to -\infty$, $\lim_{n \to \infty} \boldsymbol{X}(t_n) = \boldsymbol{Z}$. (这些术语之所以这样

称呼是因为在希腊字母表中, α 是第一个字母, 而 ω 是最后一个字母.) 下面的事实将在下一章中用到, 它的证明本质上已经包含在上面给出的证明中.

命题 一个对所有 $t \in \mathbb{R}$ 都有定义的解的 α 极限集和 ω 极限集都是闭不变集. ■

9.3 梯度系统

现在我们来考虑一类特殊类型的系统, 它们特别适合用前面的李雅普诺夫函数来研究. $\mathbb{R}^n$ 上的一个**梯度系统**是指一个形如

$$\boldsymbol{X}' = -\mathrm{grad}\, V(\boldsymbol{X})$$

的微分方程系统, 其中 $V : \mathbb{R}^n \to \mathbb{R}$ 是一个 C^∞ 函数, 且

$$\mathrm{grad}\, V = \left(\frac{\partial V}{\partial x_1}, \cdots, \frac{\partial V}{\partial x_n}\right).$$

(该系统中出现的负号是一种习惯.) 向量场 $\mathrm{grad}\, V$ 称为 V 的**梯度**. 注意, $-\mathrm{grad}\, V(\boldsymbol{X}) = \mathrm{grad}\,(-V(\boldsymbol{X}))$.

梯度系统具有一些特殊的性质, 这使得它们的流相当简单. 下面的等式是基本的:

$$DV_{\boldsymbol{X}}(\boldsymbol{Y}) = \mathrm{grad}\, V(\boldsymbol{X}) \cdot \boldsymbol{Y}.$$

也就是说, V 在 $\boldsymbol{X}$ 处的导数作用在 $\boldsymbol{Y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$ 上的值由向量 $\mathrm{grad}\, V(\boldsymbol{X})$ 和 $\boldsymbol{Y}$ 的点乘给出. 这可以由公式

$$DV_{\boldsymbol{X}}(\boldsymbol{Y}) = \sum_{j=1}^{n} \frac{\partial V}{\partial x_j}(\boldsymbol{X}) y_j$$

立刻得出.

令 $\boldsymbol{X}(t)$ 为上述梯度系统满足 $\boldsymbol{X}(0) = \boldsymbol{X}_0$ 的解, 记 $\dot{V} : \mathbb{R}^n \to \mathbb{R}$ 为 V 沿这个解的导数, 即

$$\dot{V}(\boldsymbol{X}) = \frac{\mathrm{d}}{\mathrm{d}t} V(\boldsymbol{X}(t)).$$

命题 函数 V 是系统 $\boldsymbol{X}' = -\mathrm{grad}\, V(\boldsymbol{X})$ 的一个李雅普诺夫函数. 并且, $\dot{V}(\boldsymbol{X}) = 0$ 当且仅当 $\boldsymbol{X}$ 是一个平衡点.

证明 由链式法则我们有

$$\dot{V}(\boldsymbol{X}) = DV_{\boldsymbol{X}}(\boldsymbol{X}') = \mathrm{grad}\, V(\boldsymbol{X}) \cdot (-\mathrm{grad}\, V(\boldsymbol{X})) = -|\mathrm{grad}\, V(\boldsymbol{X})|^2 \leqslant 0.$$

特别地, $\dot{V}(\boldsymbol{X}) = 0$ 当且仅当 $\mathrm{grad}\, V(\boldsymbol{X}) = 0$. ■

该命题的一个直接推论就是, 如果 $\boldsymbol{X}^*$ 是 V 的一个孤立的极小值, 则 $\boldsymbol{X}^*$ 就是上述梯度系统的一个渐近稳定的平衡点. 事实上, $\boldsymbol{X}^*$ 为孤立的就保证了在 $\boldsymbol{X}^*$ 的一个邻域中 (不包含 $\boldsymbol{X}^*$), 成立 $\dot{V} < 0$.

为了从几何上了解梯度流, 我们来看函数 $V : \mathbb{R}^n \to \mathbb{R}$ 的**水平曲面**, 也就是子集 $V^{-1}(c), c \in \mathbb{R}$. 如果 $\boldsymbol{X} \in V^{-1}(c)$ 为一个**正则点**, 即 $\operatorname{grad} V(\boldsymbol{X}) \neq 0$, 则 $V^{-1}(c)$ 在 $\boldsymbol{X}$ 附近看起来就像一个 $n-1$ 维 "曲面". 为此, 假设 (通过将坐标轴重新编号)$\partial V/\partial x_n(\boldsymbol{X}) \neq 0$, 由隐函数定理, 存在 C^∞ 函数 $g : \mathbb{R}^{n-1} \to \mathbb{R}$ 使得在 $\boldsymbol{X}$ 附近, 水平集 $V^{-1}(c)$ 由

$$x_n = g(x_1, \cdots, x_{n-1})$$

给出, 即, 在 $\boldsymbol{X}$ 附近, $V^{-1}(c)$ 看起来像函数 g 的图像. 在 $n = 2$ 的特殊情形, 当 $\boldsymbol{X}$ 是一个正则点时, $V^{-1}(c)$ 就是一条过 $\boldsymbol{X}$ 的简单曲线. 如果 $V^{-1}(c)$ 中的所有点都是正则点, 我们就称 c 是 V 的一个**正则值**. 在 $n = 2$ 的情形, 如果 c 是一个**正则值**, 则水平集 $V^{-1}(c)$ 就是一些简单 (或者, 不相交) 曲线的并. 如果 $\boldsymbol{X}$ 是 V 的一个非正则点, 则 $\operatorname{grad} V(\boldsymbol{X}) = 0$, 故 V 在 $\boldsymbol{X}$ 处的所有偏导数都是零, 从而 $\boldsymbol{X}$ 是函数 V 的一个**临界点**.

现在假设 $\boldsymbol{Y}$ 是在 $\boldsymbol{X}$ 处与水平曲面 $V^{-1}(c)$ 相切的一个向量. 则我们可在这个水平集中找到一条曲线 $\gamma(t)$ 使得 $\gamma'(0) = \boldsymbol{Y}$. 由于 V 沿 γ 为常值, 于是,

$$DV_{\boldsymbol{X}}(\boldsymbol{Y}) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} V \circ \gamma(t) = 0.$$

于是, 根据前面的观察, 我们有 $\operatorname{grad} V(\boldsymbol{X}) \cdot \boldsymbol{Y} = 0$, 或者, 换句话说, $\operatorname{grad} V(\boldsymbol{X})$ 垂直于水平集 $V^{-1}(c)$ 在 $\boldsymbol{X}$ 处的每一个切向量. 即, 在 V 的所有正则点处, 向量场 $\operatorname{grad} V(\boldsymbol{X})$ 都与水平曲面 $V^{-1}(c)$ 垂直. 我们将这些总结成下面的定理：

定理 (**梯度系统的性质**) 对于系统 $\boldsymbol{X}' = -\operatorname{grad} V(\boldsymbol{X})$,

(1) 如果 c 是 V 的一个正则值, 则该向量场垂直于水平集 $V^{-1}(c)$;

(2) V 的临界点都是该系统的平衡点;

(3) 如果临界点是 V 的一个孤立极小值, 则该点是一个渐近稳定的平衡点. ■

例 令 $V : \mathbb{R}^2 \to \mathbb{R}$ 为函数 $V(x, y) = x^2(x-1)^2 + y^2$. 则梯度系统

$$\boldsymbol{X}' = F(\boldsymbol{X}) = -\operatorname{grad} V(\boldsymbol{X})$$

由下式给出：

$$\begin{aligned} x' &= -2x(x-1)(2x-1) \\ y' &= -2y. \end{aligned}$$

这里有 3 个平衡点：$(0,0),(1/2,0),(1,0)$. 在这 3 个平衡点处线性化可得下面的矩阵：

$$DF(0,0)=\begin{pmatrix}-2&0\\0&-2\end{pmatrix},\quad DF(1/2,0)=\begin{pmatrix}1&0\\0&-2\end{pmatrix},$$

$$DF(1,0)=\begin{pmatrix}-2&0\\0&-2\end{pmatrix}.$$

于是，$(0,0)$ 和 $(1,0)$ 为汇点，而 $(1/2,0)$ 为鞍点. x 轴和 y 轴都是不变的，直线 $x=1/2$ 和 $x=1$ 也是不变的. 由于在这些竖直线上有 $y'=-2y$，从而在 $(1/2,0)$ 处的稳定曲线为直线 $x=1/2$，而在 $(1/2,0)$ 处的不稳定曲线则为 x 轴上的区间 $(0,1)$. ■

V 的水平集以及上述系统的相图如图 9.10 所示. 注意，似乎所有的解都趋于这三个平衡点中的一个. 这并不是偶然的，因为我们有：

命题 设 $\boldsymbol{Z}$ 是一个梯度流的解的 α 极限点或 ω 极限点，则 $\boldsymbol{Z}$ 为平衡点.

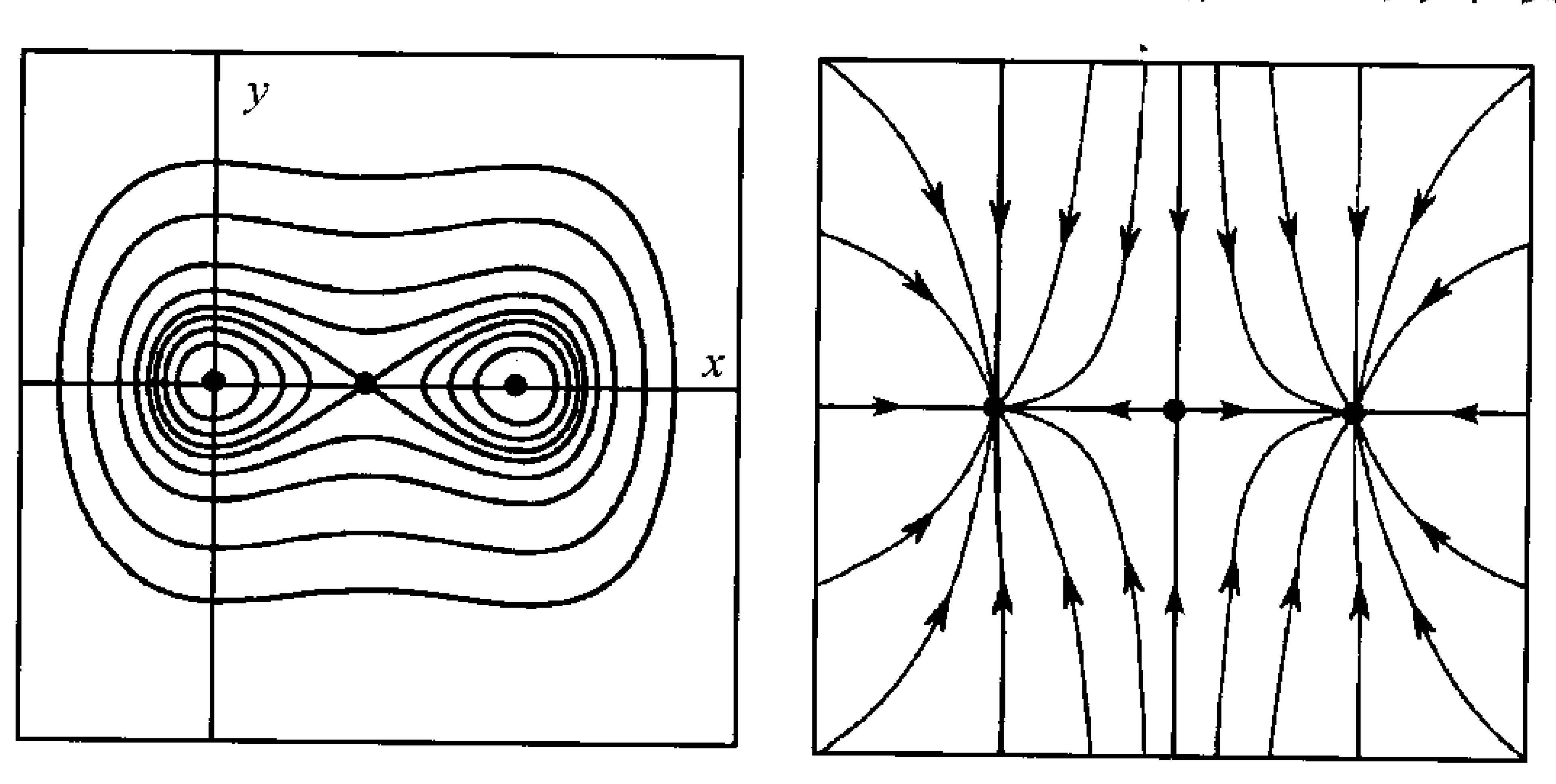


图 9.10 由 $V(x,y)=x^2(x-1)^2+y^2$ 所确定的梯度系统的水平集和相图

证明 假设 $\boldsymbol{Z}$ 是一个 ω 极限点. 像在 9.2 节中证明 Lasalle 不变原理一样，我们可以证明 V 沿过 $\boldsymbol{Z}$ 的解为常数. 从而，$\dot{V}(\boldsymbol{Z})=0$，于是 $\boldsymbol{Z}$ 必定是一个平衡点. α 极限点的情形是类似的. 事实上，$\boldsymbol{X}'=-\operatorname{grad}V(\boldsymbol{X})$ 的 α 极限点 $\boldsymbol{Z}$ 就是 $\boldsymbol{X}'=\operatorname{grad}V(\boldsymbol{X})$ 的 ω 极限点，由此可得 $\operatorname{grad}V(\boldsymbol{Z})=0$. ■

如果一个梯度系统仅有孤立的平衡点，这个结果蕴涵着，该系统的每个解必定要么趋于无穷，要么趋于一个平衡点. 在上例中，我们看到集合 $V^{-1}([0,c])$ 是有界闭的，且在梯度流下是正向不变的. 从而每个进入这种集合的解在所有 $t\geqslant 0$ 时都有定义，且趋于 3 个平衡点 $(0,0),(1/2,0),(1,0)$ 中的一个. 过每个点的解都的确进入一个这种集合，这是因为过 (x,y) 当然要进入集合 $V^{-1}([0,c_0])$，其中 $V(x,y)=c_0$.

我们来看最后一个性质，它也是所有梯度系统都共有的. 注意，在上例中，3 个平衡点处的线性化的所有特征值都是实特征值. 这也不是偶然的，因为一梯度系

统在平衡点 $\boldsymbol{X}^*$ 处的线性化是一个矩阵 $[a_{ij}]$, 其中

$$a_{ij} = -\left(\frac{\partial^2 V}{\partial x_i \partial x_j}\right)(\boldsymbol{X}^*).$$

由于混合偏导数相等, 我们有

$$\left(\frac{\partial^2 V}{\partial x_i \partial x_j}\right)(\boldsymbol{X}^*) = \left(\frac{\partial^2 V}{\partial x_j \partial x_i}\right)(\boldsymbol{X}^*),$$

于是, $a_{ij} = a_{ji}$. 由此可得, 相应的线性化系统的矩阵为一个**对称矩阵**. 我们知道, 所有这样的矩阵都只有实特征值. 例如, 在 2×2 情形, 对称矩阵都具有如下形式:

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

易见它的特征值为

$$\frac{a+c}{2} \pm \frac{\sqrt{(a-c)^2+4b^2}}{2},$$

它们都是实数. 在习题 15 中还有一个更一般点的情形. 于是, 我们有:

命题 对于梯度系统 $\boldsymbol{X}' = -\mathrm{grad}\, V(\boldsymbol{X})$, 任一平衡点处的线性化系统都只有实特征值. ■

9.4 哈密顿系统

在本节, 我们讨论另一种特殊类型的系统, 哈密顿系统. 在第 13 章中我们将看到, 从经典力学中产生的系统就是这种类型的系统.

在本节, 我们仅限于考虑 $\mathbb{R}^2$ 中的哈密顿系统. $\mathbb{R}^2$ 中的一个**哈密顿系统**是如下形式的系统:

$$\begin{aligned} x' &= \frac{\partial H}{\partial y}(x,y) \\ y' &= -\frac{\partial H}{\partial x}(x,y), \end{aligned}$$

其中 $H:\mathbb{R}^2 \to \mathbb{R}$ 是一个 C^∞ 函数, 称之为**哈密顿函数**.

例 (**无阻尼调和振子**) 回忆一下, 这个系统由下式给出:

$$\begin{aligned} x' &= y \\ y' &= -kx, \end{aligned}$$

其中 $k>0$. 这个系统的一个哈密顿函数为

$$H(x,y) = \frac{1}{2}y^2 + \frac{k}{2}x^2.$$ ■

例 (**理想单摆**) 在 9.2 节, 我们已经看到这个系统的方程为

$$\theta' = v$$
$$v' = -\sin\theta.$$

此时, 总能量函数

$$E(\theta, v) = \frac{1}{2}v^2 + 1 - \cos\theta$$

就可以作为一个哈密顿函数. 注意, 由于任何哈密顿函数加上一个常数后并不会改变方程, 所以我们总是说一个哈密顿函数.

让哈密顿系统变得如此重要的是如下的事实：哈密顿函数是一个**首次积分**, 或**运动常数**. 即, H 沿着该系统的每个解都是常数, 或者, 用前几节的语言来说就是 $\dot{H} \equiv 0$. 这可以从下面立刻得出：

$$\dot{H} = \frac{\partial H}{\partial x}x' + \frac{\partial H}{\partial y}y' = \frac{\partial H}{\partial x}\frac{\partial H}{\partial y} + \frac{\partial H}{\partial y}\left(-\frac{\partial H}{\partial x}\right) = 0.$$ ■

于是, 我们有

命题 对于 $\mathbb{R}^2$ 上的一个哈密顿系统, H 沿每个解曲线都是常数. ■

知道了一个给定系统为哈密顿的重要性在于, 此时我们不需求解系统就可以基本上作出它的相图. 假设 H 在任何开集上都不是常数, 我们先简单地作出水平曲线 $H(x, y) =$ 常数. 则系统的解就位于这些水平集上; 我们需要做的就是在这些水平集上确定解曲线的方向. 但因为我们知道了向量场, 这点很容易做到. 还要注意, 一个哈密顿系统的平衡点就出现在 H 的临界点处, 即在 H 的所有偏导数都为零的地方.

例 考虑系统

$$x' = y$$
$$y' = -x^3 + x.$$

一个哈密顿函数为

$$H(x, y) = \frac{x^4}{4} - \frac{x^2}{2} + \frac{y^2}{2} + \frac{1}{4}.$$

这里的常数 1/4 是无所谓的; 我们选它只是让 H 的极小值为 0, H 在 $(\pm 1, 0)$ 处取到这个极小值, 这些都很容易验证. 另外的一个平衡点位于原点. 该方程的线性化系统为

$$\boldsymbol{X}' = \begin{pmatrix} 0 & 1 \\ 1 - 3x^2 & 0 \end{pmatrix} \boldsymbol{X}.$$

在 $(0, 0)$ 处, 该系统的特征值为 ± 1, 从而我们得到一个鞍点. 在 $(\pm 1, 0)$ 处, 特征值为 $\pm\sqrt{2}\mathrm{i}$, 从而我们得到两个中心, 至少对线性化系统是这样的.

作出 H 的水平曲线并在非平衡点处加上方向就得到系统的相图, 如图 9.11 所示. 注意, 位于 $(\pm1, 0)$ 处的平衡点对于原来的非线性系统也仍然是中心. 还要注意, 原点处稳定曲线和不稳定曲线完全重合, 即, 我们有既正向趋于 $(0, 0)$ 又负向趋于 $(0,0)$ 的解. 这样的解就是所谓的**同宿解**或**同宿轨**. ■

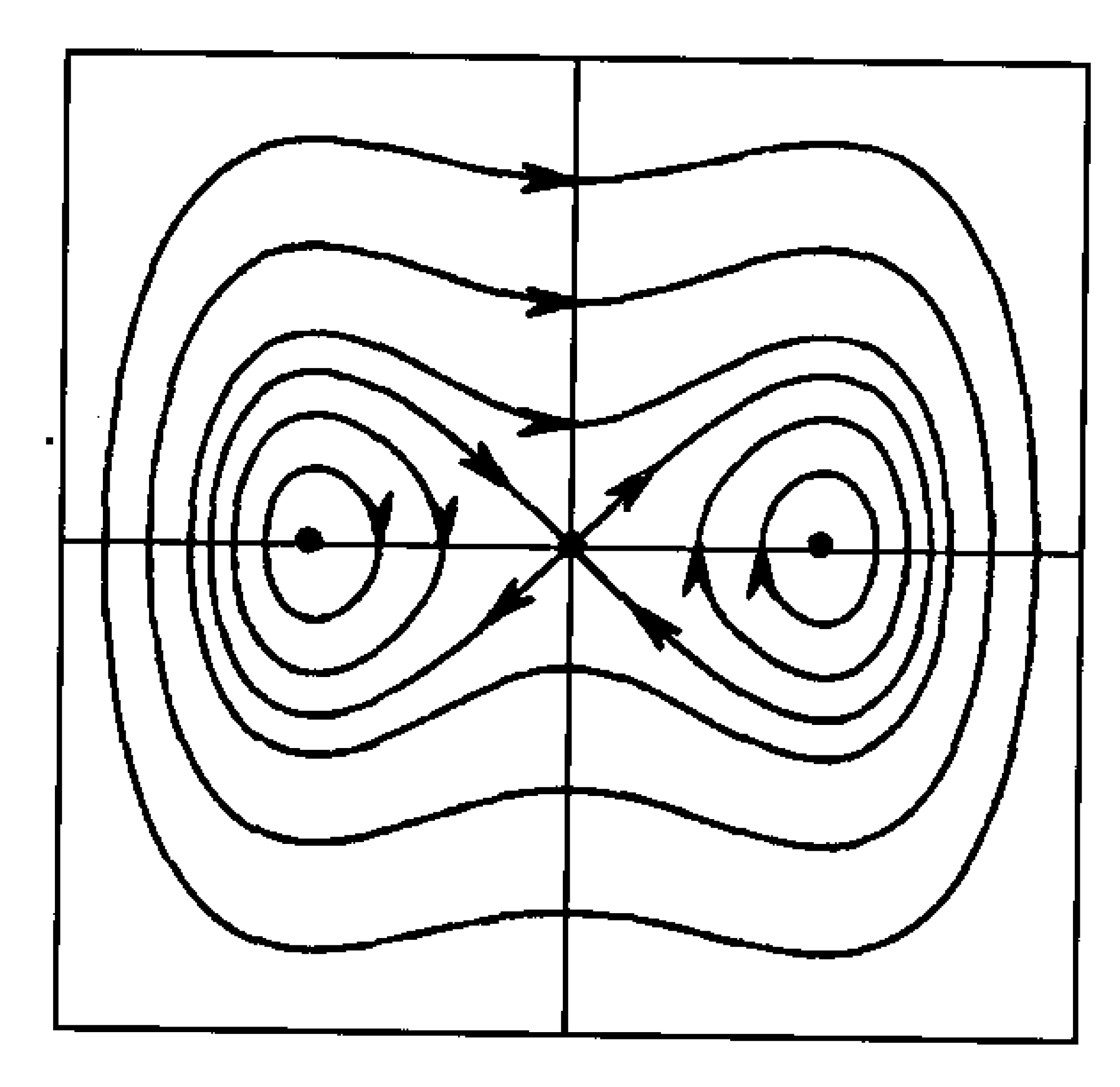


图 9.11 $x' = y,\ y' = -x^3 + x$ 的相图

该系统的特征值具有特殊的形式 ±1 和 $\pm\sqrt{2}\mathrm{i}$, 这同样不是偶然的.

命题 设 (x_0, y_0) 是一个平面哈密顿系统的平衡点. 则相应的线性化系统的特征值要么为 $\pm\lambda$, 要么为 $\pm\mathrm{i}\lambda$, 其中 $\lambda \in \mathbb{R}$. ■

该命题的证明是直接的 (见习题 11).

9.5 探索: 具有常值外力的单摆

回忆一下, 在 9.2 节中, 一个非线性单摆的方程为

$$\begin{aligned}\theta' &= v\\ v' &= -bv - \sin\theta,\end{aligned}$$

其中 θ 为单摆的角位移 (这里假设它是反时针方向测量的), 而 v 为角速度. 参数 $b > 0$ 代表阻尼的大小.

现在, 我们反时针方向对单摆施加一个常值的扭矩. 这等于在 v' 的方程上加上一个常数, 从而现在的系统就变成

$$\begin{aligned}\theta' &= v\\ v' &= -bv - \sin\theta + k,\end{aligned}$$

其中假设 $k \geqslant 0$. 由于 θ 的大小是按 $\mathrm{mod}2\pi$ 来量的, 我们可以想象该系统定义在柱面 $S^1 \times \mathbb{R}$ 上, 其中 S^1 代表单位圆周.

(1) 找出该系统的所有平衡点并确定它们的稳定性.

(2) 在 bk 参数平面上确定一些区域, 使得在不同参数区域上的系统具有不同数目的平衡点. 在每种不同的情形, 描述单摆的运动.

(3) 假设 $k > 1$. 证明该系统存在一个周期解. 提示：你能对向量场在形如 $0 < v_1 < (k - \sin\theta)/b < v_2$ 的带形区域中说些什么?

(4) 描述该系统定义在直线 $\theta = 0$ 上的庞加莱映射的定性特征.

(5) 证明在 $k > 1$ 时, 该系统具有唯一的周期解. 提示：回想能量函数

$$E(\theta, y) = \frac{1}{2}y^2 - \cos\theta + 1$$

并利用 E 沿任一周期解的总改变量必定为 0 的事实.

(6) 证明存在参数值使得稳定平衡点和周期解能同时存在.

(7) 描述当周期解不再存在时必定出现的分岔.

习　题

1. 对下面的每个系统, 作出其 x 零点集和 y 零点集, 并利用这些信息大致确定其相平面. 你可不妨假设这些系统只在 $x, y \geqslant 0$ 时有定义.
   (a) $x' = x(y + 2x - 2),\ y' = y(y - 1)$;
   (b) $x' = x(y + 2x - 2),\ y' = y(y + x - 3)$;
   (c) $x' = x(2 - y - 2x),\ y' = y(3 - 3y - x)$;
   (d) $x' = x(2 - y - 2x),\ y' = y(3 - y - 4x)$;
   (e) $x' = x(2500 - x^2 - y^2),\ y' = y(70 - y - x)$.
2. 描述系统
$$\begin{aligned} x' &= x^2 - 1 \\ y' &= -xy + a(x^2 - 1) \end{aligned}$$
   在 $a < 0$ 时的相图. 当 a 从负到正时, 流的哪些定性特征会发生变化.
3. 考虑微分方程系统
$$\begin{aligned} x' &= x(-x - y + 1) \\ y' &= y(-ax - y + b), \end{aligned}$$
   其中 $a, b > 0$, 为参数. 假设系统仅在 $x, y \geqslant 0$ 时有定义.
   (a) 对于 a, b 的不同取值, 利用零点集描述系统的相图.
   (b) 确定发生分岔时的 a, b 的取值.
   (c) 将 ab 平面分成一些区域使得在每个区域上系统有定性相似的相图, 并且描述参数穿过区域边界时的分岔.

4. 考虑系统

$$\begin{aligned} x' &= (\epsilon x + 2y)(z+1) \\ y' &= (-x + \epsilon y)(z+1) \\ z' &= -z^3. \end{aligned}$$

(a) 证明当 $\epsilon = 0$ 时, 原点不是渐近稳定的.

(b) 证明当 $\epsilon < 0$ 时, 原点的吸引盆包含区域 $z > -1$.

5. 对于非线性阻尼单摆, 证明对每个整数 n 和每个角度 θ_0, 都存在初始条件 (θ_0, v_0) 使得满足这个初始条件的解对应于单摆在趋于静止位置前, 绕着圆周转至少 n 圈, 但又不到 $n+1$ 圈.

6. 找出系统

$$\begin{aligned} x' &= -2x - y^2 \\ y' &= -y - x^2 \end{aligned}$$

的平衡点 $(0,0)$ 的一个严格李雅普诺夫函数. 找出一个尽可能大的 $\delta > 0$ 使得中心在 $(0,0)$ 半径为 δ 的开圆盘包含在 $(0,0)$ 的吸引盆中.

7. 对于下面的每个函数 $V(\boldsymbol{X})$, 简略地作出梯度系统 $\boldsymbol{X}' = -\mathrm{grad}\, V(\boldsymbol{X})$ 的相图, 在同一张图上画出 $V(\boldsymbol{X})$ 的水平曲面, 找出所有的平衡点并确定其类型.

(a) $x^2 + 2y^2$;　　(b) $x^2 - y^2 - 2x + 4y + 5$;

(c) $y \sin x$;　　(d) $2x^2 - 2xy + 5y^2 + 4x + 4y + 4$;

(e) $x^2 + y^2 - z$;　　(f) $x^2(x-1) + y^2(y-2) + z^2$.

8. 简略地作出下面系统的相图, 同时确定系统是否为哈密顿系统或梯度系统 (顺便提一下, 第一问就是一点提示).

(a) $x' = x + 2y,\ y' = -y$;　　(b) $x' = y^2 + 2xy,\ y' = x^2 + 2xy$;

(c) $x' = x^2 - 2xy,\ y' = y^2 - 2xy$;　　(d) $x' = x^2 - 2xy,\ y' = y^2 - x^2$;

(e) $x' = -\sin^2 x \sin y,\ y' = -2\sin x \cos x \cos y$.

9. 令 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$ 为一线性系统, 其中

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

(a) 确定该系统为一梯度系统时, a, b, c, d 所应满足的条件. 明确地给出一个梯度函数.

(b) 对于哈密顿系统重做上一问.

10. 考虑平面系统

$$\begin{aligned} x' &= f(x,y) \\ y' &= g(x,y). \end{aligned}$$

明确地给出该系统为梯度系统或哈密顿系统时 f, g 应满足的条件.

11. 证明平面哈密顿系统的平衡点处的线性化的特征值要么是 $\pm\lambda$ 要么是 $\pm i\lambda$, 其中 $\lambda \in \mathbb{R}$.
12. 令 T 是将正方形 $0 \leqslant \theta_1, \theta_2 \leqslant 2\pi$ 的对边等同得到的环面. 设 $F(\theta_1, \theta_2) = \cos\theta_1 + \cos\theta_2$. 在 T 上简略地作出系统 $-\operatorname{grad} F$ 的相图. 当把 T 表示成一个油炸圈饼的表面时, 简略地作出相图的一个三维表示.
13. 假设上题中的 F 为一哈密顿函数时, 重做上题.
14. 在环面 T(如上题) 上, 令 $F(\theta_1, \theta_2) = \cos\theta_1(2 - \cos\theta_2)$. 在 T 上简略地作出系统 $-\operatorname{grad} F$ 的相图. 当把 T 表示成一个油炸圈饼的表面时, 简略地作出相图的一个三维表示.
15. 证明 3×3 的对称矩阵只有实特征值.
16. 一个系统的解 $\boldsymbol{X}(t)$ 称为**回复的**, 如果对于某个序列 $t_n \to \infty$, 有 $\boldsymbol{X}(t_n) \to \boldsymbol{X}(0)$. 证明梯度系统没有非常值的回复解.
17. 证明有界闭的 ω 极限集是连通的. 举出一个平面系统的例子, 使得它的 ω 极限集由两条平行线构成.

第 10 章　闭轨和极限集

在前面几章, 我们集中在微分方程系统的平衡解上. 虽然平衡解无疑是最重要的解, 但是一些其他类型的解也同样重要. 在本章, 我们将研究另一种重要类型的解, **周期解**, 或称**闭轨**. 回忆一下, 在系统 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 中出现周期解, 如果我们有一个非平衡点 $\boldsymbol{X}$ 以及时间 $\tau>0$ 使得 $\phi_\tau(\boldsymbol{X})=\boldsymbol{X}$. 此时, 对所有的 t, 都有 $\phi_{t+\tau}(\boldsymbol{X})=\phi_t(\boldsymbol{X})$, 从而 ϕ_t 是一个周期函数. 最小的这样的 $\tau>0$ 称为这个解的**周期**. 例如, 无阻尼调和振子方程的所有非零解都是周期解. 类似渐近稳定的平衡点, 周期解也可能吸引其他的解. 即, 解可以以周期解为极限, 这就像解可以趋于平衡点一样.

在平面上, 解的极限行为本质上就限制在一些平衡点和闭轨上, 尽管也会有一些例外情形. 在本章, 我们将通过重要的庞加莱 - 本迪克逊定理来研究这一现象. 在后面我们将看到, 在高于二位的情形, 解的极限行为可以比现在的要复杂得多的多.

10.1　极　限　集

我们首先描述微分方程系统的解的极限行为. 回忆一下, $\boldsymbol{Y}\in\mathbb{R}^n$ 是过 $\boldsymbol{X}$ 的解的一个 ω 极限点, 如果存在序列 $t_n\to\infty$ 使得 $\lim_{n\to\infty}\phi_{t_n}(\boldsymbol{X})=\boldsymbol{Y}$. 即, 当时间向前时, 过 $\boldsymbol{X}$ 的解曲线在 $\boldsymbol{Y}$ 点聚集. 过 $\boldsymbol{X}$ 的解的所有 ω 极限点的集合就是 $\boldsymbol{X}$ 的 ω 极限集, 记为 $\omega(\boldsymbol{X})$. 在上面的定义中, 把 $t_n\to\infty$ 换成 $t_n\to-\infty$, 就得到 α 极限点和 α 极限集 $\alpha(\boldsymbol{X})$. $\omega(\boldsymbol{X})$ 和 $\alpha(\boldsymbol{X})$ 统称为**极限集**.

下面是极限集的几个例子. 如果 $\boldsymbol{X}^*$ 是一个渐近平衡点, 则它是它的吸引盆中的每个点的 ω 极限集. 任何平衡点既是它自己的 α 极限集又是它自己的 ω 极限集. 一个周期解是它上面的每个点的 α 极限集和 ω 极限集. 这样一个解也可能是许多其他点的 ω 极限集.

例　考虑用极坐标给出的平面系统

$$r'=\frac{1}{2}(r-r^3)$$

$$\theta'=1.$$

在 8.1 节中我们已经看到, 该方程的所有非零解都趋于位于平面上单位圆周的周

期解 (见图 10.1). 从而, 任何非零解的 ω 极限集就是这个闭轨. ■

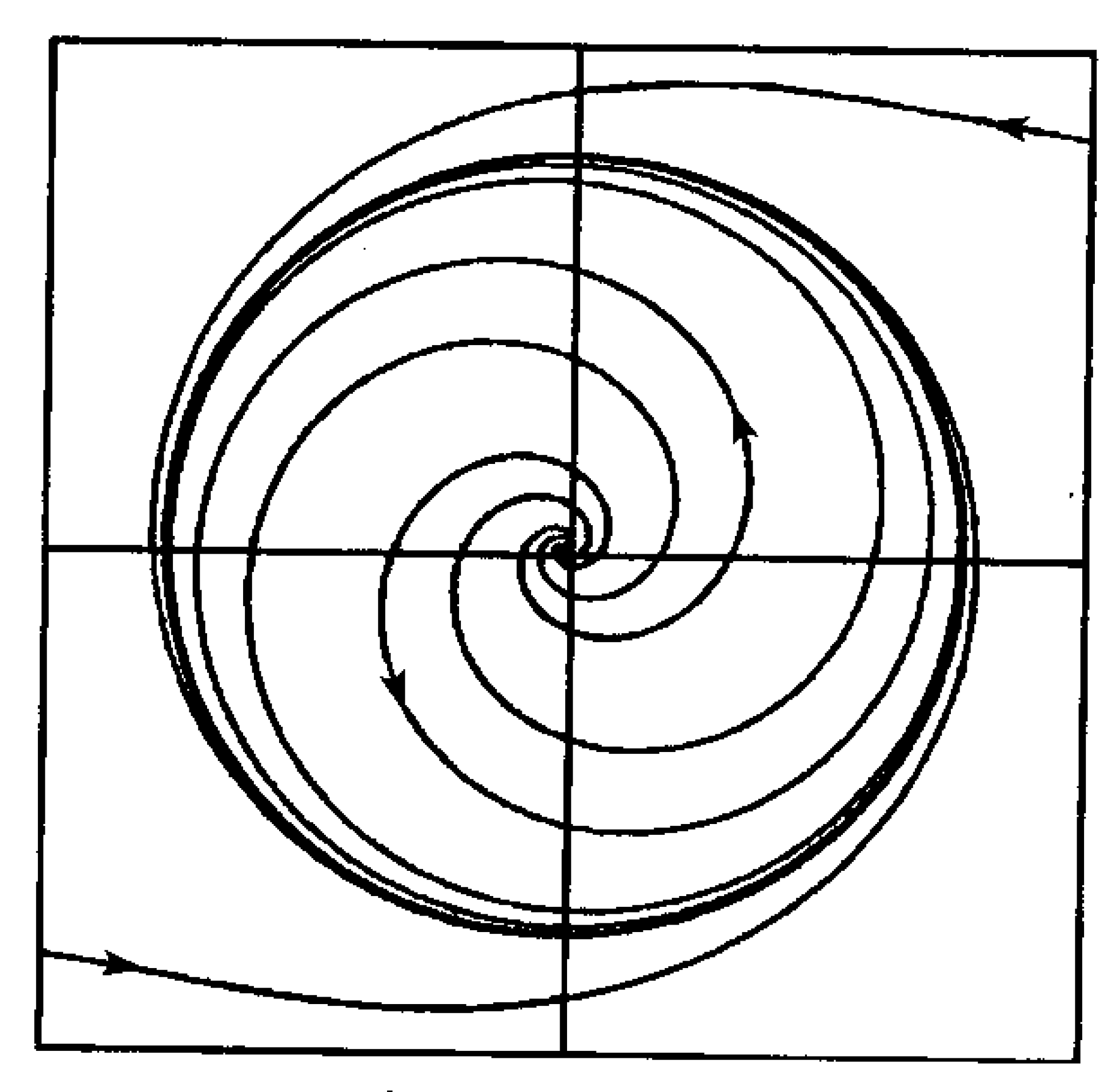


图 10.1 $r' = \frac{1}{2}(r - r^3)$, $\theta' = 1$ 的相平面

例 考虑系统

$$x' = \sin x(-0.1\cos x - \cos y)$$
$$y' = \sin y(\cos x - 0.1\cos y).$$

在正方形的四个角 $(0,0)$, $(0,\pi)$, (π,π), $(\pi,0)$ 上, 有 4 个是鞍点的平衡点, 同时还有许多其他的平衡点. 有一些异宿解按照上面所列的顺序将这 4 个平衡点连接起来 (见图 10.2). 在 $(\pi/2,\pi/2)$ 处还有一个螺线源点. 所有从这个源点散发出来的解聚集在连接这些平衡点的 4 个异宿解上 (见本章的习题 4). 于是, 这些解上的任何点的 ω 极限点就是由 $x = 0,\pi$ 和 $y = 0,\pi$ 围成的正方形. ■

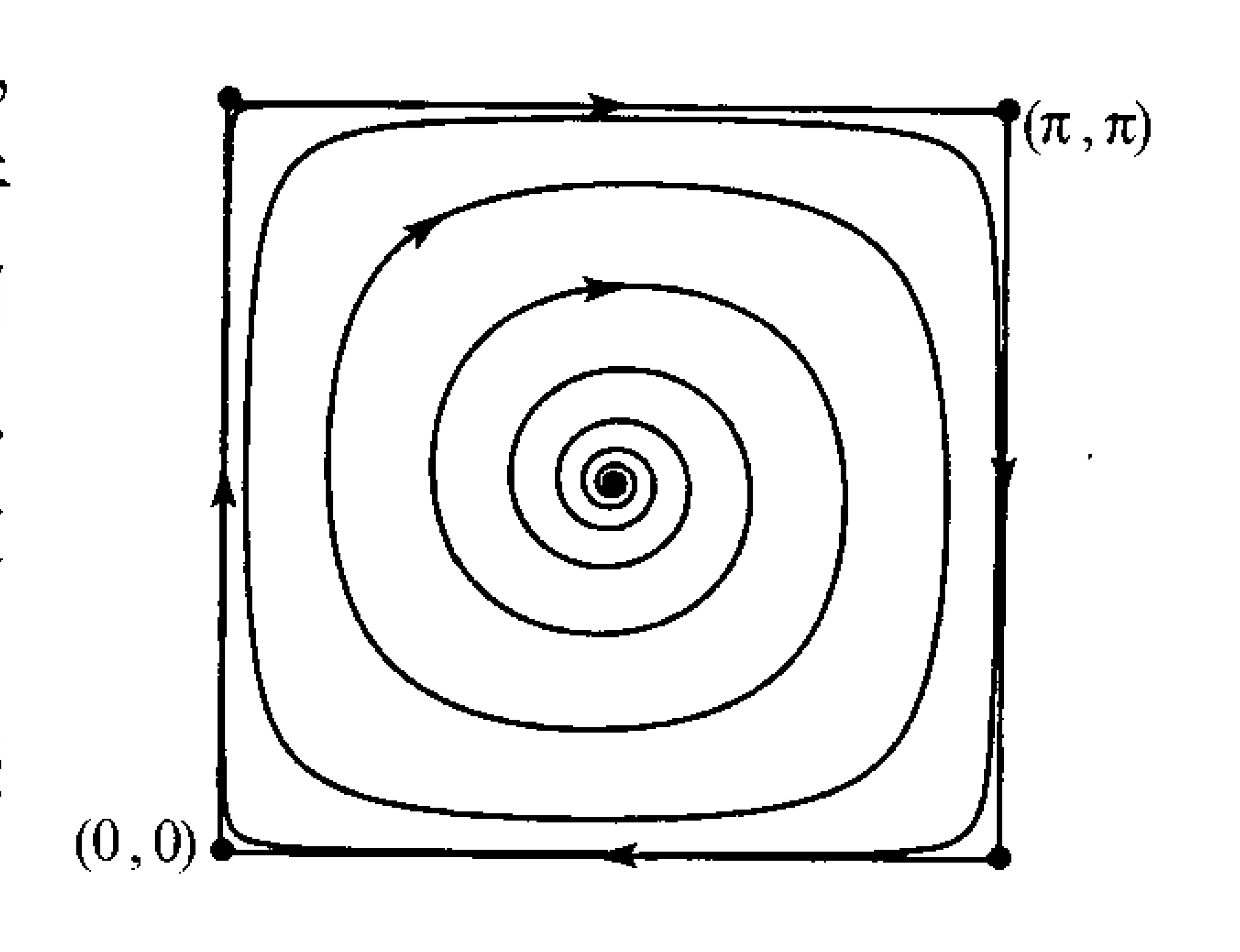


图 10.2 任何从位于 $(\pi/2,\pi/2)$ 处的源点散发出来的解的 ω 极限集为由 4 个平衡点以及 4 个异宿解围成的正方形

在三维时, 有一些及其复杂的极限集的例子, 描述它们不是很容易. 然而, 在平面情形, 极限集则相当简单. 事实上, 图 10.2 是典型的, 因为可以证明, 异于闭轨或平衡点的闭有界极限集就是由一些平衡点以及一些连接它们的解构成. 将在 10.5 节中讨论的庞加莱 - 本迪克逊定理断言, 平面上不含平衡点的闭有界极限集必定是一个闭轨.

从 9.2 节还知道, 一个极限集在 $\mathbb{R}^n$ 中闭, 而且在流的作用下不变. 我们将来还需要用到下面的结果:

命题

(1) 如果 $\boldsymbol{X}$ 和 $\boldsymbol{Z}$ 位于同一条解曲线上, 则 $\omega(\boldsymbol{X}) = \omega(\boldsymbol{Z})$, $\alpha(\boldsymbol{X}) = \alpha(\boldsymbol{Z})$;

(2) 如果 D 是一个正向不变的闭集, $\boldsymbol{Z}\in D$, 则 $\omega(\boldsymbol{Z})\subset D$; 对于负向不变集和 α 极限集有类似的结果.

(3) 一个闭不变集, 特别地, 一个极限集, 包含它里面的每个点的 α 极限集和 ω 极限集.

证明　对于 (1), 假设 $\boldsymbol{Y}\in\omega(\boldsymbol{X})$, 且 $\phi_s(\boldsymbol{X})=\boldsymbol{Z}$. 如果 $\phi_{t_n}(\boldsymbol{X})\to\boldsymbol{Y}$, 则我们有

$$\phi_{t_n-s}(\boldsymbol{Z})=\phi_{t_n}(\boldsymbol{X})\to\boldsymbol{Y}.$$

从而, 也有 $\boldsymbol{Y}\in\omega(\boldsymbol{Z})$. 对于 (2), 如果当 $t_n\to\infty$ 时, $\phi_{t_n}(\boldsymbol{Z})\to\boldsymbol{Y}\in\omega(\boldsymbol{Z})$, 则我们有, 对于充分大的 n 有 $t_n>0$, 故 $\phi_{t_n}(\boldsymbol{Z})\in D$. 由于 D 为闭集, 于是就有 $\boldsymbol{Y}\in D$. 最后, (3) 可从 (2) 立刻推出. ■

10.2　局部截面和流盒

在本章剩下的部分, 我们将讨论限制在平面系统. 在本节, 我们描述 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的流在一个非平衡点的给定点 $\boldsymbol{X}_0$ 附近的局部行为. 我们的目的是先在 $\boldsymbol{X}_0$ 处构造一个局部截面, 然后在 $\boldsymbol{X}_0$ 的一个邻域中构造一个流盒. 在这个流盒中, 系统的解行为特别地简单.

假设 $\boldsymbol{F}(\boldsymbol{X}_0)\neq 0$. $\boldsymbol{X}_0$ 处的**横截线**就是过 $\boldsymbol{X}_0$ 点且垂直于基点在 $\boldsymbol{X}_0$ 的向量 $\boldsymbol{F}(\boldsymbol{X}_0)$ 的直线, 将它记为 $l(\boldsymbol{X}_0)$. 我们将 $l(\boldsymbol{X}_0)$ 如下参数化. 令 $\boldsymbol{V}_0$ 是一个基点在 $\boldsymbol{X}_0$ 且垂直于 $\boldsymbol{F}(\boldsymbol{X}_0)$ 的单位向量. 然后由 $h(u)=\boldsymbol{X}_0+u\boldsymbol{V}_0$ 定义 $h:\mathbb{R}\to l(\boldsymbol{X}_0)$.

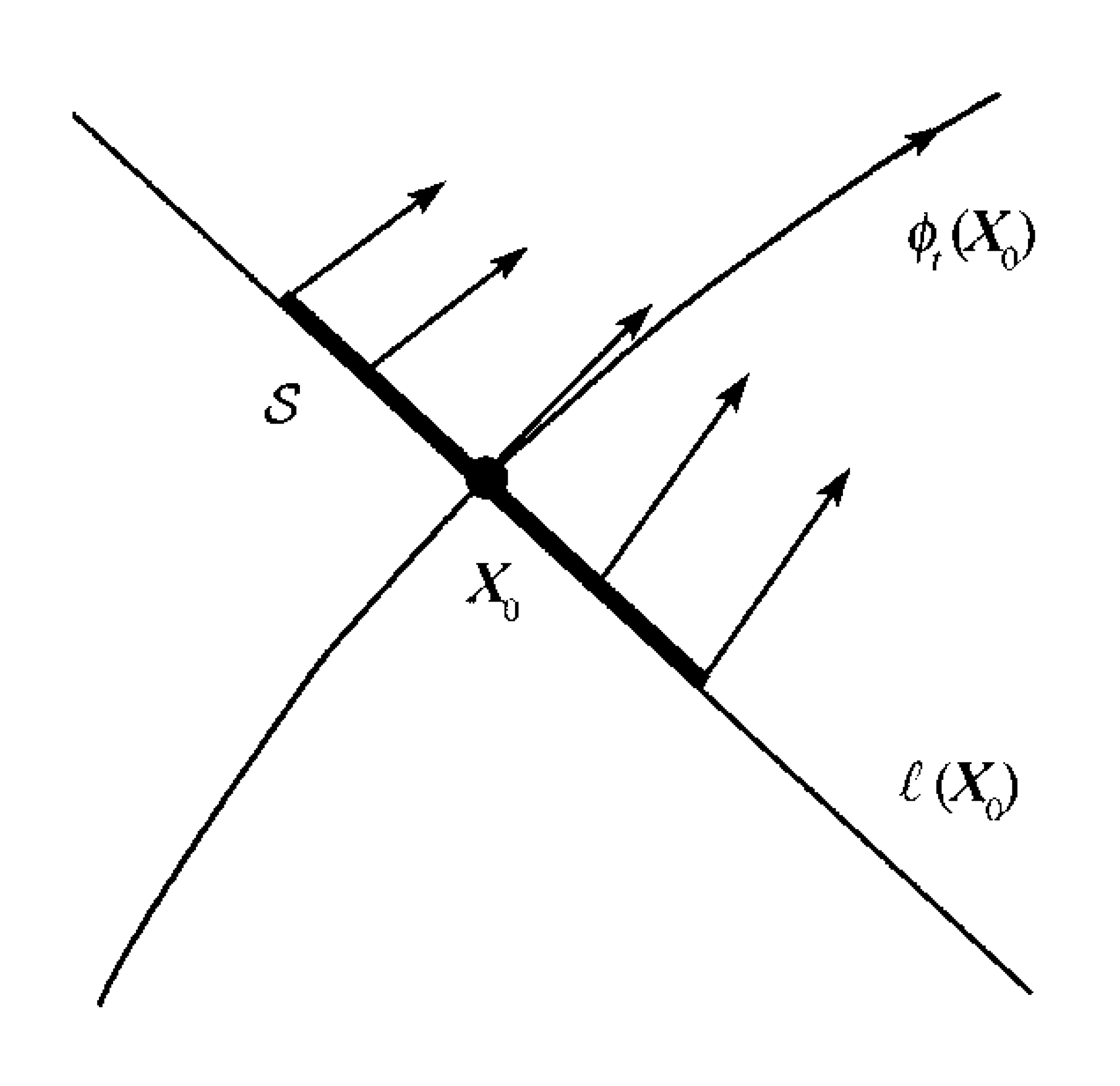


图 10.3　X_0 处的一个局部截面 $\mathcal{S}$ 和向量场沿 $\mathcal{S}$ 的几个代表向量

由于 $\boldsymbol{F}(\boldsymbol{X})$ 是连续的, 至少在 $l(\boldsymbol{X}_0)$ 上一个包含 $\boldsymbol{X}_0$ 的开区间内, 这个向量场与 $l(\boldsymbol{X}_0)$ 不相切. 我们称这样的一个包含 $\boldsymbol{X}_0$ 的开区间为 $\boldsymbol{X}_0$ 处的一个**局部截面**. 在一个局部截面 $\mathcal{S}$ 上的每一点, 向量场总是指向"离开"$\mathcal{S}$, 从而解总是穿过一个局部截面. 特别地, 对于 $\boldsymbol{X}\in\mathcal{S}$, $\boldsymbol{F}(\boldsymbol{X})\neq\boldsymbol{0}$(见图 10.3).

$\boldsymbol{X}_0$ 处的局部截面的第一个应用就是在 $\boldsymbol{X}_0$ 的一个邻域中构造相应的流盒. 一个流盒就是通过一组特殊的坐标, 在一个非平衡点的邻域, 给出流行为的一个完全的描述. 流在流盒中的直观描述很简单：所有的点都沿平行线匀速运动.

给定 $\boldsymbol{X}_0$ 处的一个局部截面 $\mathcal{S}$, 可以如下构造一个从 $\mathbb{R}^2$ 的原点的一个邻域 $\mathcal{N}$ 到 $\boldsymbol{X}_0$ 的一个邻域的映射 Ψ. 给定 $(s,u)\in\mathbb{R}^2$, 我们定义

$$\Psi(s,u)=\phi_s(h(u)),$$

其中 h 为如上所述的横截线的参数化. 注意, Ψ 将 $\mathcal{N}$ 中的竖直线 $(0,u)$ 映到局部截面 $\mathcal{S}$; 而且 Ψ 将 $\mathcal{N}$ 中的水平线映到系统的各条解曲线. 如果我们选取 $\mathcal{N}$ 充分小, 则映射 Ψ 在 $\mathcal{N}$ 上是一一的. 还需注意, $D\Psi$ 将 $\mathcal{N}$ 中的常值向量场 $(1,0)$ 映到向量场 $\boldsymbol{F}(\boldsymbol{X})$. 用第 4 章的语言来说就是, Ψ 就是常值向量场的流与非线性系统的流之间的一个局部共轭.

通常将 $\mathcal{N}$ 取为形式 $\{(s,u)||s|<\sigma\}$, 其中 $\sigma>0$. 此时, 我们常常记 $\mathcal{V}_\sigma=\Psi(\mathcal{N})$, 并称 $\mathcal{V}_\sigma$ 是 $\boldsymbol{X}_0$ 处 (或关于 $\boldsymbol{X}_0$) 的**流盒**(见图 10.4). 流盒的一个重要性质就是, 如果 $\boldsymbol{X}\in\mathcal{V}_\sigma$, 则存在一个唯一的 $t\in(-\sigma,\sigma)$ 使得 $\phi_t(\boldsymbol{X})\in\mathcal{S}$.

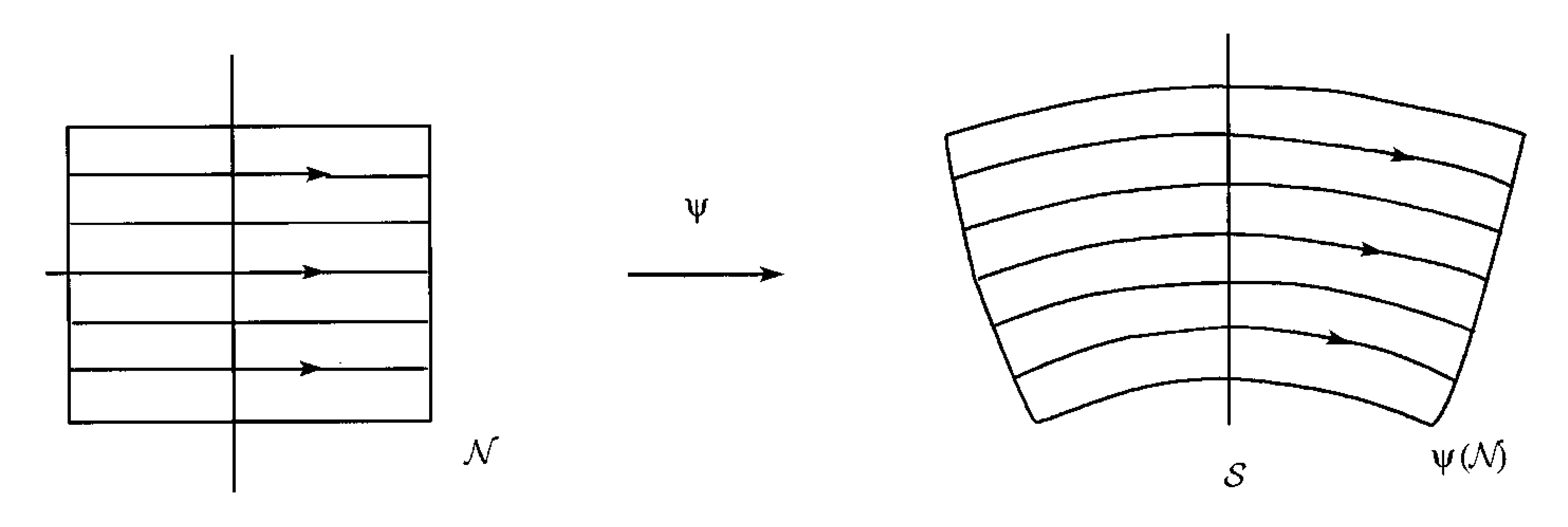


图 10.4 属于 $\mathcal{S}$ 的流盒

如果 $\mathcal{S}$ 是一个局部截面, 则过一点 $\boldsymbol{Z}_0$(可能离 $\mathcal{S}$ 很远) 的解可能在某个时刻 t_0 达到 $\boldsymbol{X}_0\in\mathcal{S}$(见图 10.5). 我们来证明, 在某种局部意义下, 首次到达 $\mathcal{S}$ 的时间是 $\boldsymbol{Z}_0$ 的一个连续函数. 更准确地说就是:

命题 令 $\mathcal{S}$ 为 $\boldsymbol{X}_0$ 处的一个局部截面, 且 $\phi_{t_0}(\boldsymbol{Z}_0)=\boldsymbol{X}_0$. 令 $\mathcal{W}$ 为 $\boldsymbol{Z}_0$ 的一个邻域. 则, 存在一个包含 $\boldsymbol{Z}_0$ 的开集 $\mathcal{U}\subset\mathcal{W}$ 以及一个可微函数 $\tau:\mathcal{U}\to\mathbb{R}$ 满足 $\tau(\boldsymbol{Z}_0)=t_0$, 且对每个 $\boldsymbol{X}\in\mathcal{U}$ 都有

$$\phi_{\tau(\boldsymbol{X})}(\boldsymbol{X})\in\mathcal{S}.$$

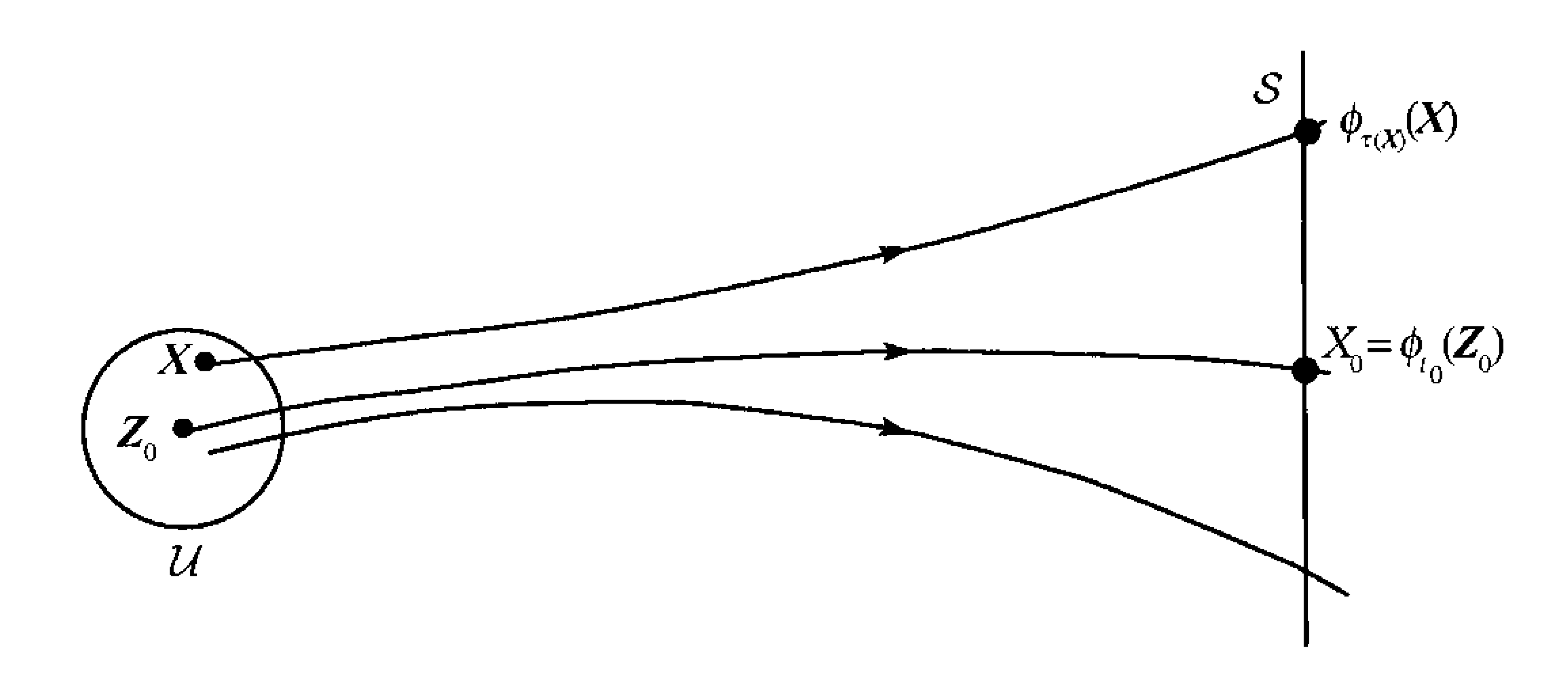


图 10.5 解穿过局部截面 $\mathcal{S}$

证明　假设 $\boldsymbol{F}(\boldsymbol{X}_0)$ 为向量 (α,β), 注意, $(\alpha,\beta)\neq(0,0)$. 对 $\boldsymbol{Y}=(y_1,y_2)\in\mathbb{R}^2$, 定义 $\eta:\mathbb{R}^2\to\mathbb{R}$ 如下:

$$\eta(\boldsymbol{Y})=\boldsymbol{Y}\cdot\boldsymbol{F}(\boldsymbol{X}_0)=\alpha y_1+\beta y_2.$$

注意, $\boldsymbol{Y}$ 属于横截线 $l(\boldsymbol{X}_0)$ 当且仅当 $\boldsymbol{Y}=\boldsymbol{X}_0+\boldsymbol{V}$, 其中 $\boldsymbol{V}\cdot\boldsymbol{F}(\boldsymbol{X}_0)=0$. 于是, $\boldsymbol{Y}\in l(\boldsymbol{X}_0)$ 当且仅当 $\eta(\boldsymbol{Y})=\boldsymbol{Y}\cdot\boldsymbol{F}(\boldsymbol{X}_0)=\boldsymbol{X}_0\cdot\boldsymbol{F}(\boldsymbol{X}_0)$.

现在定义 $G:\mathbb{R}^2\times\mathbb{R}\to\mathbb{R}$ 如下:

$$G(\boldsymbol{X},t)=\eta(\phi_t(\boldsymbol{X}))=\phi_t(\boldsymbol{X})\cdot\boldsymbol{F}(\boldsymbol{X}_0).$$

由于 $\phi_{t_0}(\boldsymbol{X}_0)=\boldsymbol{Z}_0$, 我们有 $G(\boldsymbol{Z}_0,t_0)=\boldsymbol{X}_0\cdot\boldsymbol{F}(\boldsymbol{X}_0)$. 而且还有

$$\frac{\partial G}{\partial t}(\boldsymbol{Z}_0,t_0)=|\boldsymbol{F}(\boldsymbol{X}_0)|^2\neq 0.$$

然后, 由隐函数定理可知, 存在定义在 $\boldsymbol{Z}_0$ 的某个邻域 $\mathcal{U}_1$ 的函数 $\tau:\mathbb{R}^2\to\mathbb{R}$ 满足 $\tau(\boldsymbol{Z}_0)=t_0$ 以及

$$G(\boldsymbol{X},\tau(\boldsymbol{X}))\equiv G(\boldsymbol{Z}_0,t_0)=\boldsymbol{X}_0\cdot F(\boldsymbol{X}_0).$$

从而 $\phi_{\tau(\boldsymbol{X})}(\boldsymbol{X})$ 属于横截线 $l(\boldsymbol{X}_0)$. 如果 $\mathcal{U}\subset\mathcal{U}_1$ 为 $\boldsymbol{Z}_0$ 的一个充分小的邻域, 则 $\phi_{t(\boldsymbol{X})}(\boldsymbol{X})\in\mathcal{S}$, 这正是我们要证明的. ■

10.3　庞加莱映射

和平衡点情形类似, 闭轨也可能为稳定的、渐近稳定的或不稳定的. 这些概念对闭轨的定义完全类似于 8.4 节中对平衡点的相应情形. 然而, 确定闭轨的稳定性则要比对平衡点做相应的问题要困难得多. 我们的确需要一个与确定大多数平衡点稳定性的线性化技巧相类似的工具, 虽然这个工具往往在实际中非常难使用. 下面来介绍一个这样的工具.

给定一个闭轨 γ, 有一个与之相应的**庞加莱映射**. 我们已经在前面的 1.4 节和 6.2 节中遇到过几个这样的例子. 在一个闭轨附近, 这个映射可以如下定义. 取一点 $\boldsymbol{X}_0\in\gamma$, 令 $\mathcal{S}$ 为 $\boldsymbol{X}_0$ 处的一个局部截面. 我们来考虑 $\mathcal{S}$ 的首次返回映射. 这就是函数 P, 它将 $\boldsymbol{X}\in\mathcal{S}$ 映到点 $P(\boldsymbol{X})=\phi_t(\boldsymbol{X})$, 其中 t 是满足 $\phi_t(\boldsymbol{X})\in\mathcal{S}$ 的最小正时间. 由于经过 $\mathcal{S}$ 上的某些点的解可能永远也不返回 $\mathcal{S}$, 因而, P 不一定在 $\mathcal{S}$ 上的所有点处都有定义. 但是, 我们一定有 $P(\boldsymbol{X}_0)=\boldsymbol{X}_0$, 而且上一个命题保证了 P 在 $\boldsymbol{X}_0$ 的一个邻域上是有定义的且是连续可微的.

在平面系统情形, 一个局部截面就是一条过 $\boldsymbol{X}_0$ 的直线的子集, 因而我们可将这时的局部截面看成是 $\mathbb{R}$ 的一个子集并且取 $\boldsymbol{X}_0=0\in\mathbb{R}$. 于是, 庞加莱映射就是

一个将 0 映到 0 的实函数. 如果 $|P'(0)| < 1$, 则 $P(x)$ 具有形式 $P(x) = ax+$ 高阶项, 其中 $|a| < 1$. 于是, 对于 0 附近的 x, $P(x)$ 就比 x 更靠近 0. 这意味着, 经过局部截面 $\mathcal{S}$ 上的相应点的解在穿过 $\mathcal{S}$ 一次后将变得更接近 γ. 再这样继续下去, 我们就可看出, 每穿过 $\mathcal{S}$ 一次都把解带得更加接近 γ, 由此可得, γ 为渐近稳定的. 我们有:

命题 令 $\boldsymbol{X}' = F(\boldsymbol{X})$ 为一平面系统, 假设 $\boldsymbol{X}_0$ 位于一条闭轨 γ 上. 令 P 是定义在某个局部截面上的 $\boldsymbol{X}_0$ 的一个邻域上的庞加莱映射. 如果 $|P'(\boldsymbol{X}_0)| < 1$, 则 γ 是渐近稳定的. ■

例 考虑用极坐标给出的平面系统

$$\begin{aligned} r' &= r(1-r) \\ \theta' &= 1. \end{aligned}$$

显然, 有一条位于单位圆周 $r = 1$ 上的闭轨. 在直角坐标系中, 当初始条件为 $(1, 0)$ 时, 这个周期解由 $(\cos t, \sin t)$ 给出. 又由于 $\theta' = 1$, 存在一个沿正实轴方向的局部截面. 进一步, 任给 $x \in (0, \infty)$, $\phi_{2\pi}(x, 0)$ 也在正实轴 $\mathbb{R}^+$ 上. 于是, 我们得到一个庞加莱映射 $P : \mathbb{R}^+ \to \mathbb{R}^+$. 又由于 $x = 1, y = 0$ 为周期解的初始条件, 因而 $P(1) = 1$. 为了验证这个解的稳定性, 我们需要计算 $P'(1)$.

为此, 我们来计算从 $(x, 0)$ 出发的解. 我们有 $\theta(t) = t$, 还需找出 $r(2\pi)$. 为了计算 $r(t)$, 分离变量后我们得到

$$\int \frac{\mathrm{d}r}{r(1-r)} = t + \text{常数}.$$

积分之后得到

$$r(t) = \frac{x\mathrm{e}^t}{1 - x + x\mathrm{e}^t}.$$

于是,

$$P(x) = r(2\pi) = \frac{x\mathrm{e}^{2\pi}}{1 - x + x\mathrm{e}^{2\pi}}.$$

求导可得 $P'(2\pi) = 1/\mathrm{e}^{2\pi}$, 从而就有 $0 < P'(1) < 1$. 于是, 周期解是渐近稳定的. ■

机敏的读者可能已经注意到了这里的故事. 为了确定庞加莱映射, 实际上我们首先找到了所有从 $(x, 0)$ 出发的解. 那么, 我们为什么需要计算庞加莱映射? 不错, 问得好! 的确, 计算庞加莱映射的准确表达式常常是非常困难的, 即使是它沿一个周期轨道的导数也是如此, 这是因为, 实际中我们很少有得到闭轨的表达式的时候, 就更不用说附近的解了. 我们将看到, 在建立一个特定系统的几何模型时 (见第 14 章的洛伦茨系统), 庞加莱映射往往会比较有用. 在有些情形, 我们可以克服这个困难从而深入了解庞加莱映射, 例如在 12.3 节中研究范德波尔方程时就是如此.

10.4 平面动力系统中的单调序列

令 $\boldsymbol{X}_0, \boldsymbol{X}_1, \cdots \in \mathbb{R}^2$ 为过 $\boldsymbol{X}_0$ 的解曲线上的不同点的有限或无限序列. 我们称这个序列**沿这个解是单调的**, 如果存在 $0 \leqslant t_1 < t_2 < \cdots$, 使得 $\phi_{t_n}(\boldsymbol{X}_0) = \boldsymbol{X}_n$.

令 $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \cdots$ 是位于 $\mathbb{R}^2$ 中一线段 I 上的有限或无限点列. 我们称这个点列**在 I 上是单调的**, 如果对所有的 $n \geqslant 1$, 在 I 上的自然序下, $\boldsymbol{Y}_n$ 总在 $\boldsymbol{Y}_{n-1}$ 和 $\boldsymbol{Y}_{n+1}$ 之间.

一个点列可能在一条解曲线和线段 I 的交上; 它们可能沿该解曲线单调而沿线段不单调, 或者反过来 (见图 10.6). 但是, 如果该线段在平面上是一个局部截面, 这种情况是不可能出现的.

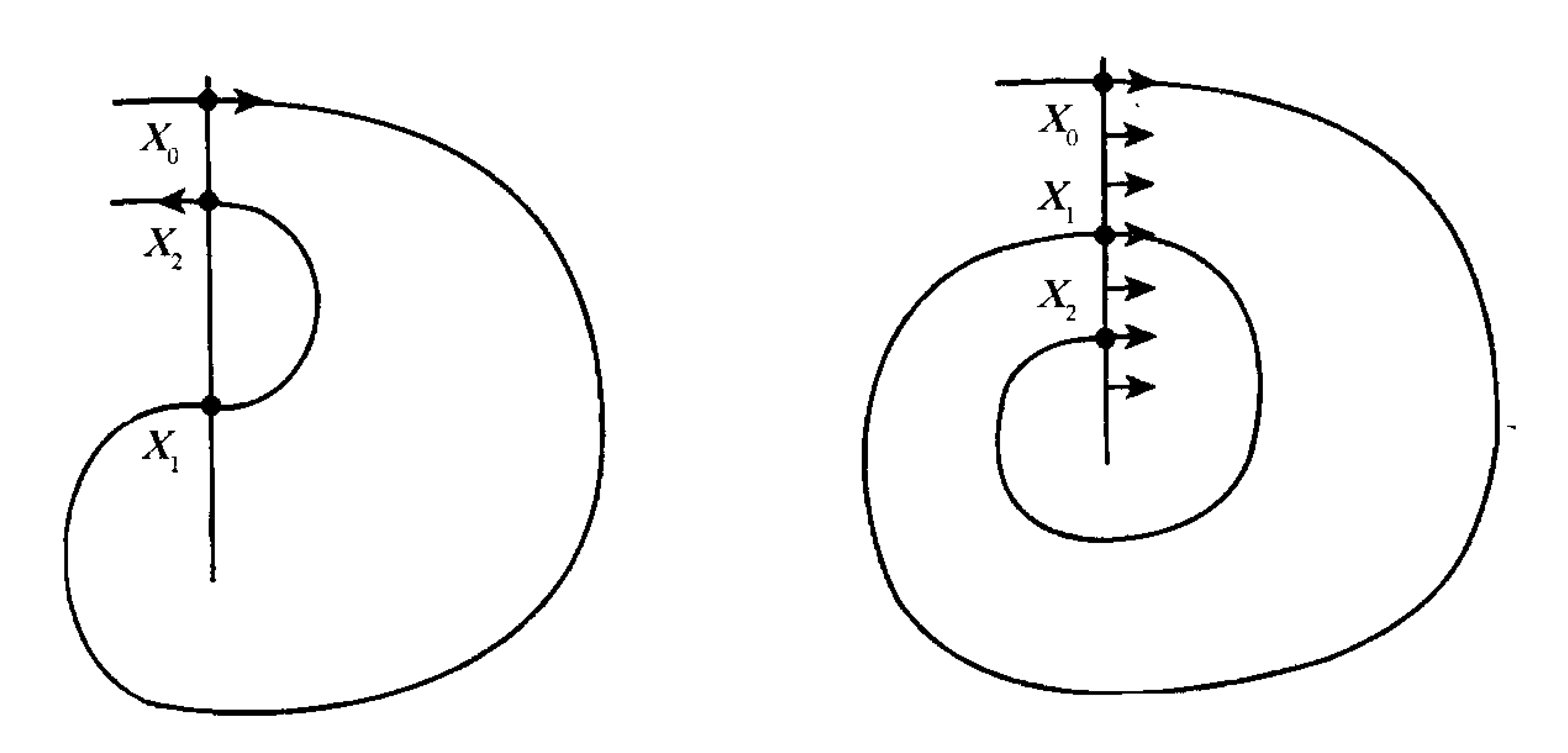


图 10.6 两个解穿过一条直线. 左边, $\boldsymbol{X}_0, \boldsymbol{X}_1, \boldsymbol{X}_2$ 沿解是单调的但沿直线则不是. 右边, $\boldsymbol{X}_0, \boldsymbol{X}_1, \boldsymbol{X}_2$ 沿解和直线都是单调的

命题 令 $\mathcal{S}$ 是平面微分方程系统的一个局部截面, $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots$ 是 $\mathcal{S}$ 上位于同一解曲线上的一列不同的点. 如果该序列沿这个解是单调的, 则它沿 $\mathcal{S}$ 也是单调的.

证明 只需考虑 $\mathcal{S}$ 上的 3 个点 $\boldsymbol{Y}_0, \boldsymbol{Y}_1, \boldsymbol{Y}_2$ 就可以了. 令 Σ 是由这个解在 $\boldsymbol{Y}_0, \boldsymbol{Y}_1$ 之间的部分以及 $\boldsymbol{Y}_0, \boldsymbol{Y}_1$ 之间的线段 $T \subset \mathcal{S}$ 构成的简单闭曲线. 记 D 是由 Σ 围成的区域. 我们假设穿过 $\boldsymbol{Y}_1$ 的解在 $\boldsymbol{Y}_1$ 处离开 D(见图 10.7; 如果解进入 D, 讨论是类似的). 由于 T 是局部截面的一部分, 因而解在 T 中的每个点都离开 D.

由此可得, D 的余集是正向不变的. 这是因为, 没有解可以从 T 上的点进入 D; 又由解的唯一性, 解也不能穿过连接 $\boldsymbol{Y}_0$ 和 $\boldsymbol{Y}_1$ 的解曲线.

从而, 对所有的 $t > 0$, $\phi_t(\boldsymbol{Y}_1) \in \mathbb{R}^2 - D$. 特别地, $\boldsymbol{Y}_2 \in \mathcal{S} - T$. 集合 $\mathcal{S} - T$ 为两个半开区间 I_0 和 I_1 的并, 且 $\boldsymbol{Y}_j$ 为 I_j $(j = 0, 1)$ 的一个端点. 可以从 $\phi_\epsilon(\boldsymbol{Y}_1)$(这里 $\epsilon > 0$ 非常小) 作一个不穿过 Σ 的弧段到 $\boldsymbol{Y}_1$. 因而, I_1 位于 D 的外部. 同理可

得, I_0 位于 D 的内部. 由于 $\boldsymbol{Y}_2$ 必定在 D 的外部, 故 $\boldsymbol{Y}_2 \in I_1$. 这就证明了, 在 I 上, $\boldsymbol{Y}_1$ 在 $\boldsymbol{Y}_0$ 和 $\boldsymbol{Y}_2$ 之间, 这也就证明了这个命题. ■

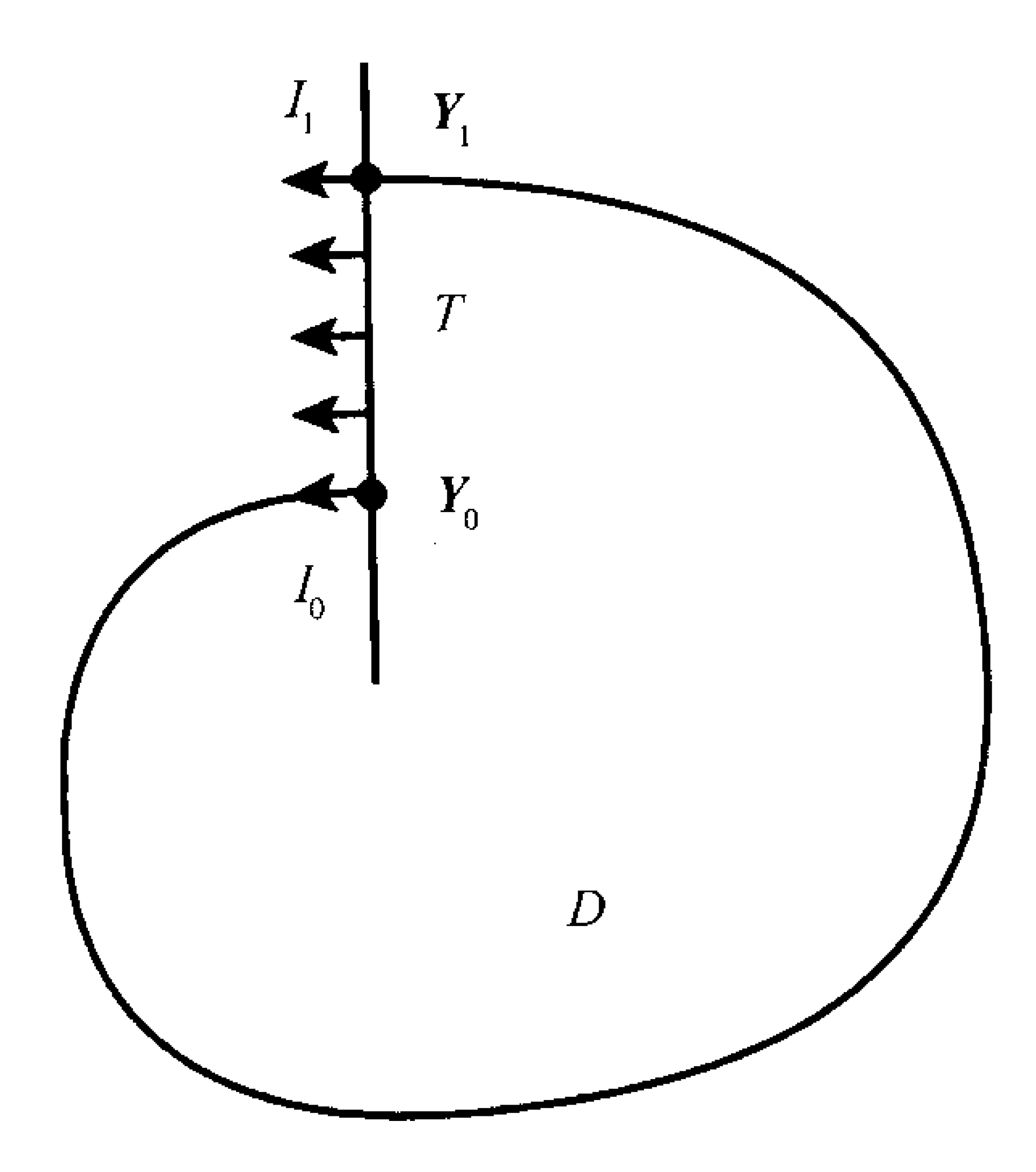


图 10.7 解穿过 T 离开区域 D

现在我们来看极限点的一个重要性质.

命题 假设, 对一个平面系统, $\boldsymbol{Y} \in \omega(\boldsymbol{X})$. 则, 对于任何局部截面, 过 $\boldsymbol{Y}$ 的解至多穿过一次. 同样的结论对 $\boldsymbol{Y} \in \alpha(\boldsymbol{X})$ 也成立.

证明 为了得到矛盾, 假设 $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ 是过 $\boldsymbol{Y}$ 的解上的两个不同的点, 而且 $\mathcal{S}$ 是包含 $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ 的一个局部截面. 假设 $\boldsymbol{Y} \in \omega(\boldsymbol{X})$ (对 $\alpha(\boldsymbol{X})$ 的讨论是类似的). 于是, $\boldsymbol{Y}_k \in \omega(\boldsymbol{X}), k = 1, 2$. 令 $\mathcal{V}_k$ 是由某一区间 $J_k \subset \mathcal{S}$ 所定义的 $\boldsymbol{Y}_k$ 处的流盒; 我们假设 J_1 和 J_2 不相交, 如图 10.8 所示. 过 $\boldsymbol{X}$ 的解将无穷次进入 $\mathcal{V}_k$, 从而要无穷次穿过 J_k. 于是就有一个沿过 $\boldsymbol{X}$ 的解单调的序列

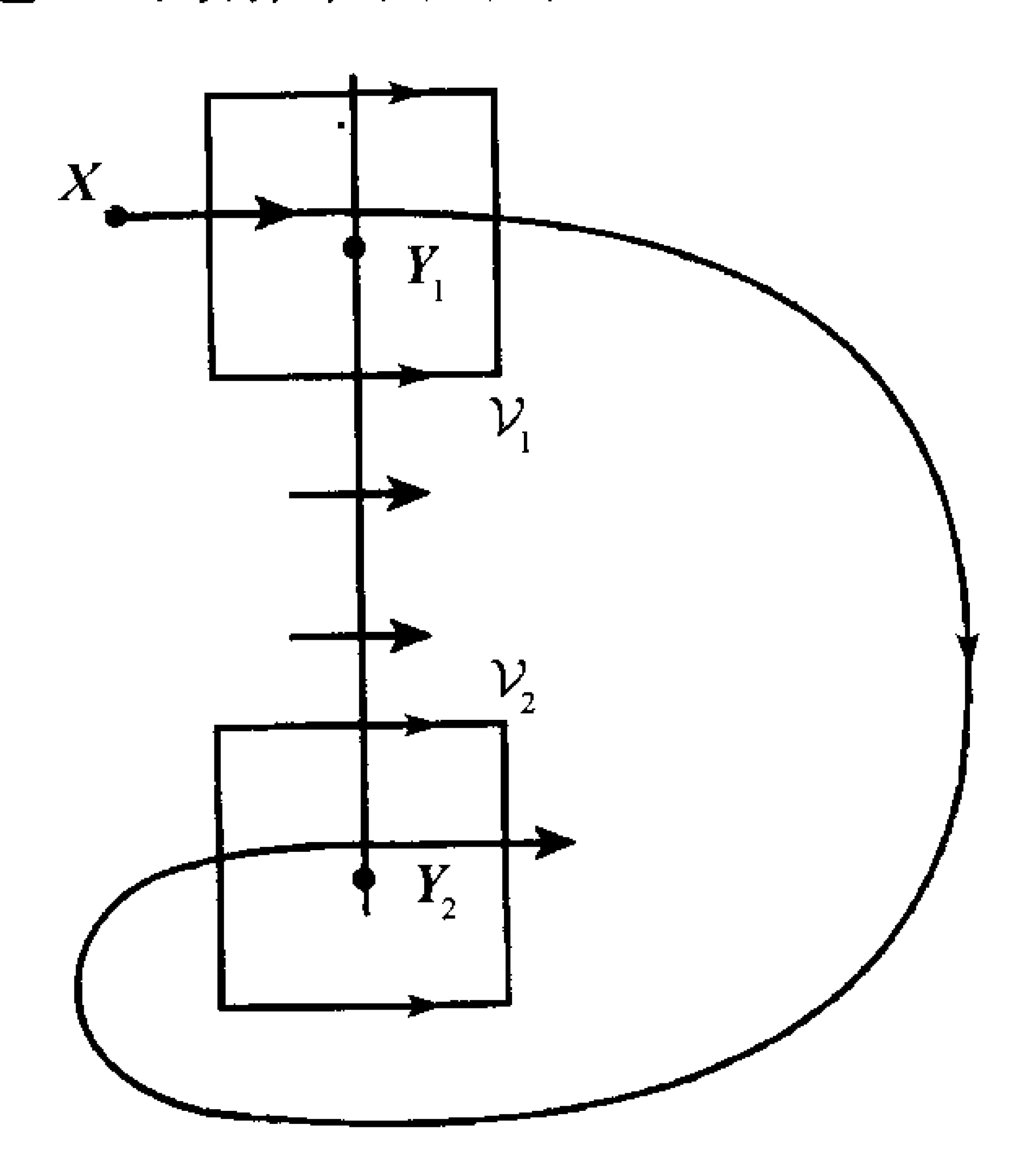


图 10.8 过 $\boldsymbol{X}$ 的解不可能无穷次穿过 $\mathcal{V}_1$ 和 $\mathcal{V}_2$

$$a_1, b_1, a_2, b_2, a_3, b_3, \cdots,$$

其中 $a_n \in J_1, b_n \in J_2, n = 1, 2, \cdots$. 由于 J_1 和 J_2 不相交, 这样的一个序列不可能沿 $\mathcal{S}$ 是单调的, 这与上一个命题的结论矛盾. ■

10.5 庞加莱 - 本迪克逊定理

在本节, 我们证明关于平面系统的著名结果:

定理 **(庞加莱 - 本迪克逊)** 假设 Ω 是平面微分方程系统的一个非空的、闭的、有界的极限集, 且它不包含平衡点, 则 Ω 是一条闭轨.

证明 假设 $\omega(\boldsymbol{X})$ 是闭有界的, 且 $\boldsymbol{Y} \in \omega(\boldsymbol{X})$. ($\alpha$ 极限集的情形是类似的.) 我们先证明 $\boldsymbol{Y}$ 位于一条闭轨上, 然后再证明这条闭轨就是 $\omega(\boldsymbol{X})$.

由于 $\boldsymbol{Y}$ 属于 $\omega(\boldsymbol{X})$, 从 10.1 节可知, $\omega(\boldsymbol{Y})$ 是 $\omega(\boldsymbol{X})$ 的一个非空子集. 令 $\boldsymbol{Z} \in \omega(\boldsymbol{Y})$, $\mathcal{S}$ 是 $\boldsymbol{Z}$ 处的一个局部截面. 再令 $\mathcal{V}$ 是相应于 $\mathcal{S}$ 的流盒. 根据上一节的结果可知, 过 $\boldsymbol{Y}$ 的解与 $\mathcal{S}$ 正好交于一点. 另一方面, 存在序列 $t_n \to \infty$ 使得 $\phi_{t_n}(\boldsymbol{Y}) \to \boldsymbol{Z}$; 从而有无穷多个 $\phi_{t_n}(\boldsymbol{Y})$ 属于 $\mathcal{V}$. 因而, 我们可以找到 $r, s \in \mathbb{R}$ 使得 $r > s$ 且 $\phi_r(\boldsymbol{Y}), \phi_s(\boldsymbol{Y}) \in \mathcal{S}$. 于是 $\phi_r(\boldsymbol{Y}) = \phi_s(\boldsymbol{Y})$, 即 $\phi_{r-s}(\boldsymbol{Y}) = \boldsymbol{Y}$ 且 $r - s > 0$. 由于 $\omega(\boldsymbol{X})$ 不含平衡点, 因而 $\boldsymbol{Y}$ 必定位于一个闭轨上.

剩下需要证明, 如果 γ 是 $\omega(\boldsymbol{X})$ 中的一条闭轨, 则 $\gamma = \omega(\boldsymbol{X})$. 为此, 只需证明

$$\lim_{t \to \infty} d(\phi_t(\boldsymbol{X}), \gamma) = 0,$$

其中 $d(\phi_t(\boldsymbol{X}), \gamma)$ 为 $\phi_t(\boldsymbol{X})$ 到集合 γ 的距离, 也就是 $\phi_t(\boldsymbol{X})$ 到 γ 的最近点之间的距离.

令 $\mathcal{S}$ 为某点 $\boldsymbol{Y} \in \gamma$ 处的一个局部截面. 令 $\epsilon > 0$, 考虑相应于 $\mathcal{S}$ 的流盒 $\mathcal{V}_\epsilon$. 则存在序列 $t_0 < t_1 < \cdots$ 满足

(1) $\phi_{t_n}(\boldsymbol{X}) \in \mathcal{S}$;

(2) $\phi_{t_n}(\boldsymbol{X}) \to Y$;

(3) $\phi_t(\boldsymbol{X}) \notin \mathcal{S}$, 其中 $t_{n-1} < t < t_n$, $n = 1, 2, \cdots$.

记 $\boldsymbol{X}_n = \phi_{t_n}(\boldsymbol{X})$. 根据上一节的第一个命题, $\boldsymbol{X}_n$ 是 $\mathcal{S}$ 上的一个单调序列, 且收敛到 $\boldsymbol{Y}$.

我们断言, 对于充分大的 n, 正数 $t_{n+1} - t_n$ 构成的集合有一个上界. 为了看出这一点, 假设 $\phi_\tau(\boldsymbol{Y}) = \boldsymbol{Y}$, 其中 $\tau > 0$. 则当 $\boldsymbol{X}_n$ 充分接近 $\boldsymbol{Y}$ 时, 有 $\phi_\tau(\boldsymbol{X}_n) \in \mathcal{V}_\epsilon$, 于是对某个 $t \in [-\epsilon, \epsilon]$,

$$\phi_{\tau+t}(\boldsymbol{X}_n) \in \mathcal{S}.$$

从而,

$$t_{n+1} - t_n \leqslant \tau + \epsilon.$$

这就给出了 $t_{n+1} - t_n$ 的一个上界. 又, 显然 $t_{n+1} - t_n$ 至少是 2ϵ, 从而当 $n \to \infty$ 时, $t_n \to \infty$.

取一个很小的 $\beta > 0$. 根据解对初始条件的连续性可知, 存在 $\delta > 0$ 使得如果 $|\boldsymbol{Z} - \boldsymbol{Y}| < \delta$ 且 $|t| \leqslant \tau + \epsilon$, 则 $|\phi_t(\boldsymbol{Z}) - \phi_t(\boldsymbol{Y})| < \beta$. 即, 对于所有满足 $|t| \leqslant \tau + \epsilon$ 的 t, 解 $\phi_t(\boldsymbol{Z})$ 到 γ 的距离都小于 β. 取 n_0 充分大使得当 $n \geqslant n_0$ 时, $|\boldsymbol{X}_n - \boldsymbol{Y}| < \delta$. 于是, 当 $|t| \leqslant \tau + \epsilon$ 且 $n \geqslant n_0$ 时, 有

$$|\phi_t(\boldsymbol{X}_n) - \phi_t(\boldsymbol{Y})| < \beta.$$

现在令 $t \geqslant t_{n_0}$. 取 $n \geqslant n_0$ 满足

$$t_n \leqslant t \leqslant t_{n+1}.$$

于是,

$$\begin{aligned} d(\phi_t(\boldsymbol{X}), \gamma) &\leqslant |\phi_t(\boldsymbol{X}) - \phi_{t-t_n}(\boldsymbol{Y})| \\ &= |\phi_{t-t_n}(\boldsymbol{X}_n) - \phi_{t-t_n}(\boldsymbol{Y})| \\ &< \beta, \end{aligned}$$

其中最后一个不等式是因为 $|t - t_n| \leqslant \tau + \epsilon$. 这表明对所有充分大的 t, $\phi_t(\boldsymbol{X})$ 到 γ 的距离都小于 β. 这就完成了庞加莱-本迪克逊定理的证明. ■

例　同宿解 是另一种既不是闭轨又不是平衡点的 ω 极限集的例子. 考虑系统

$$x' = -y - \left(\frac{x^4}{4} - \frac{x^2}{2} + \frac{y^2}{2}\right)(x^3 - x)$$

$$y' = x^3 - x - \left(\frac{x^4}{4} - \frac{x^2}{2} + \frac{y^2}{2}\right)y.$$

计算表明系统有 3 个平衡点：$(0,0), (-1,0), (1,0)$. 原点是一个鞍点, 而其他两个平衡点都是源点. 该系统的相图如图 10.9 所示. 注意, 原点有一对同宿解, 即它们从原点出发而最终又回到原点; 而且远离原点的解都趋向于聚集在原点和这对同宿解上; 而从每个源点散发出来的解的 ω 极限集都只由一个同宿解以及原点构成. 这些事实的证明见习题 6. ■

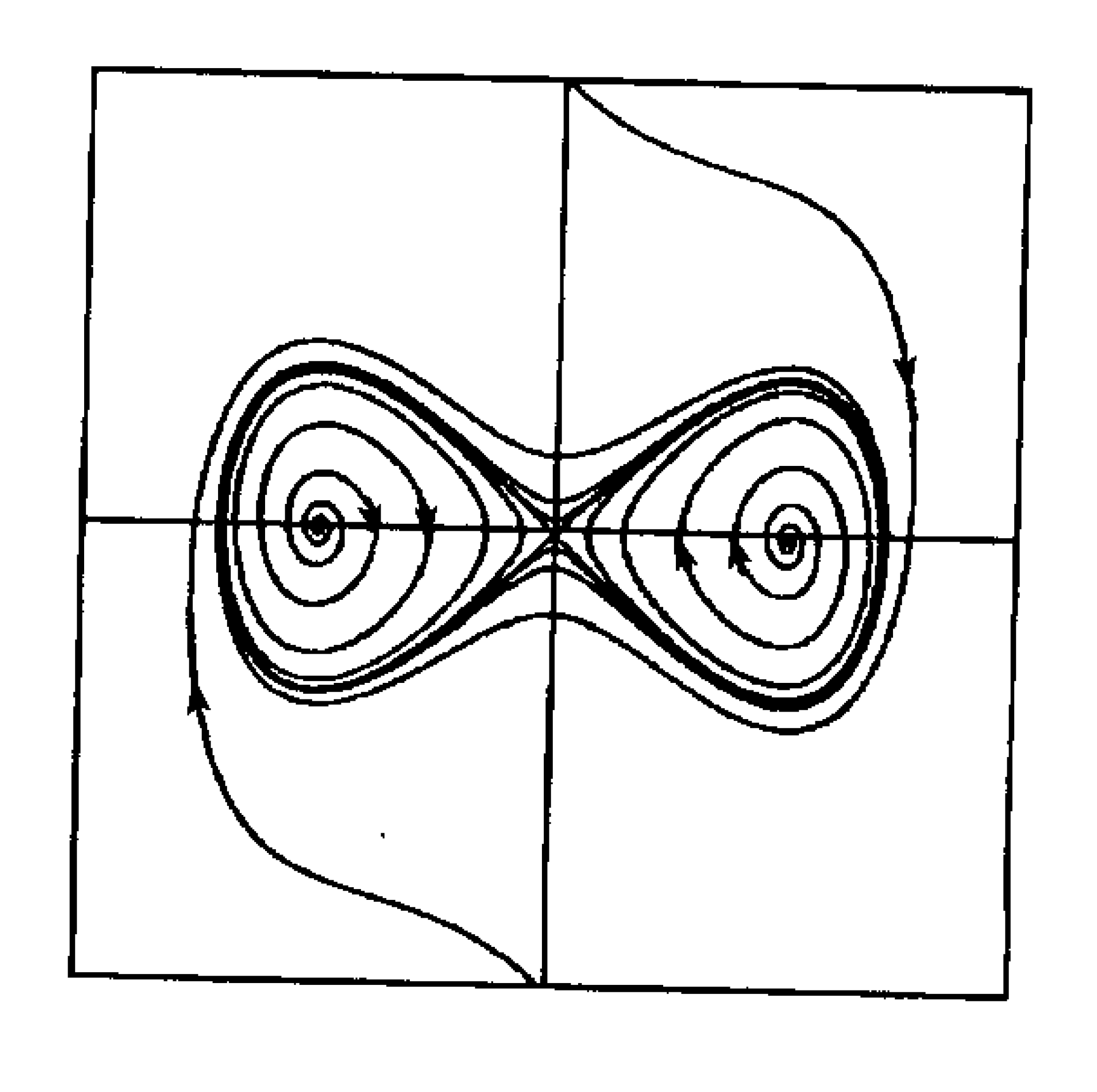


图 10.9　ω 极限集中的一对同宿解

10.6 庞加莱-本迪克逊定理的应用

庞加莱-本迪克逊定理从本质上决定了平面流的所有可能的极限行为. 我们将在本节给出这个重要定理的很多推论.

如果一个闭轨 γ 满足, 存在 $\boldsymbol{X} \notin \gamma$ 使得 $\gamma \subset \omega(\boldsymbol{X})$ 或 $\gamma \subset \alpha(\boldsymbol{X})$, 则称 γ 为一个**极限环**. 在第一种情形, γ 称为一个 ω 极限环; 在第二种情形, γ 称为一个 α 极限环. 在本节我们只处理 ω 极限集的情形; 通过简单地将时间反向就可处理 α 极限集的情形.

在庞加莱-本迪克逊定理的证明中, 我们已经得出极限环具有如下的性质: 如果 γ 是一个 ω 极限环, 则存在 $\boldsymbol{X} \notin \gamma$ 使得

$$\lim_{t\to\infty} d(\phi_t(\boldsymbol{X}), \gamma) = 0.$$

几何上, 这意味着某个解在 $t \to \infty$ 时盘旋地趋于 γ (见图 10.10). 并不是所有的闭轨都具有这个性质. 例如, 如果一个线性系统在 $\mathbb{R}^2$ 的原点是一个中心, 则就没有解逼近那些包围原点的闭轨, 从而它们都不是极限环.

极限环都具有一种 (至少是单边的) 稳定性. 令 γ 为一个极限环, 且假设在 $t \to \infty$ 时, $\phi_t(\boldsymbol{X})$ 盘旋地趋于 γ. 令 $\mathcal{S}$ 是在 $\boldsymbol{Z} \in \gamma$ 的一个局部截面. 则存在与 γ 不相交的一个区间 $T \subset \mathcal{S}$, 它的两个端点分别为 $\phi_{t_0}(\boldsymbol{X}), \phi_{t_1}(\boldsymbol{X})$, 其中 $t_0 < t_1$, 而且它与过 $\boldsymbol{X}$ 的解在 $t_0 < t < t_1$ 时没有交点. 见图 10.11. 其中的环形区域 A 的一个边界为 γ, 另外一个边界由 T 与曲线

$$\{\phi_t(\boldsymbol{X}) | t_0 \leqslant t \leqslant t_1\}$$

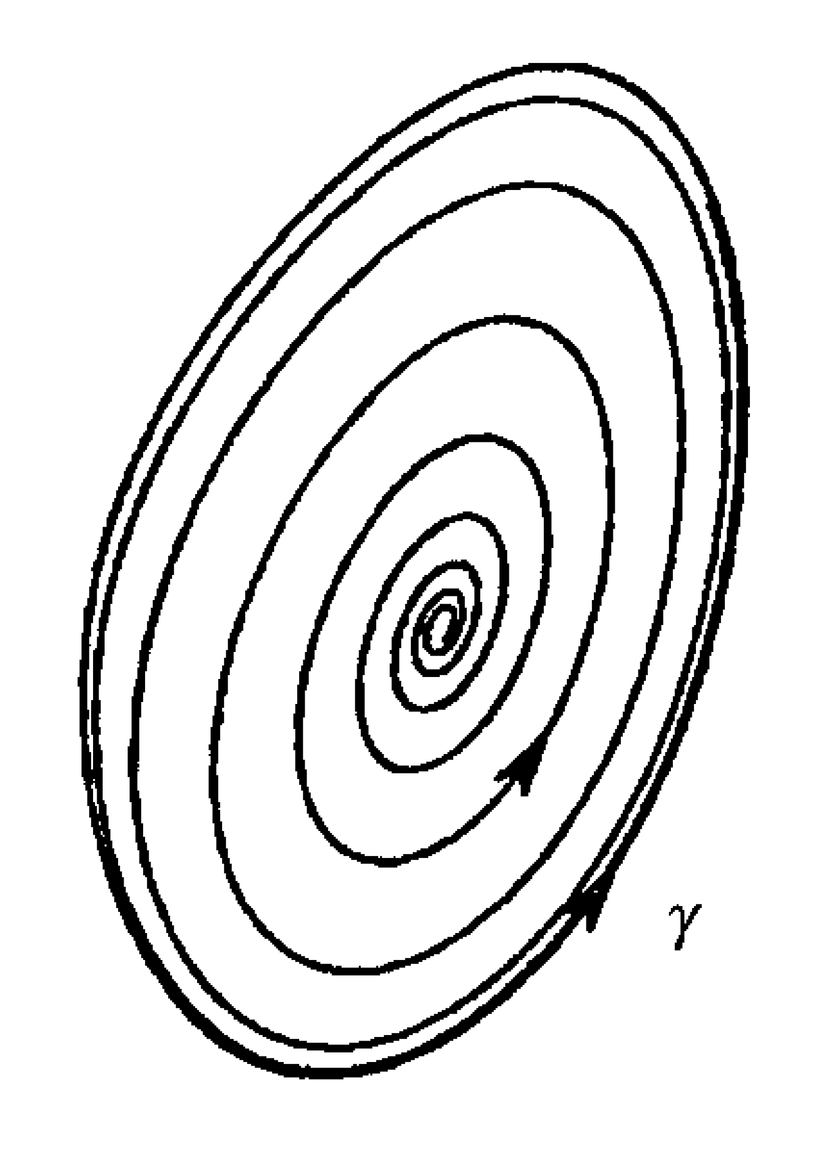


图 10.10 一个解盘旋地趋向于一个极限环

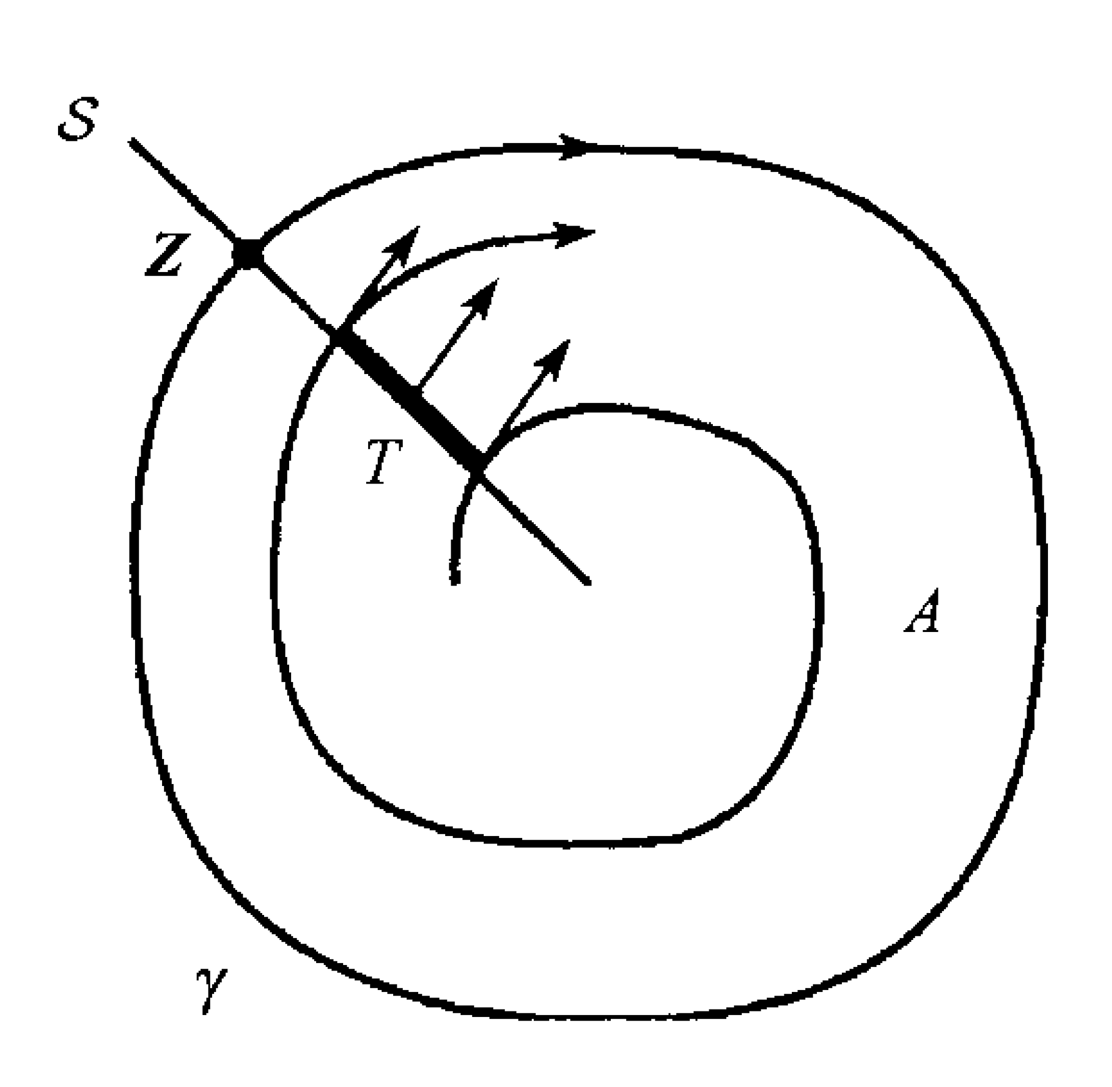


图 10.11 区域 A 是正向不变的

的并构成; 可见 A 以及 $B = A - \gamma$ 都是正向不变的. 容易看出, 对于所有的 $\boldsymbol{Y} \in B$, $\phi_t(\boldsymbol{Y})$ 都盘旋地趋于 γ. 这样我们就有:

推论 1 令 γ 为一个 ω 极限环. 如果 $\gamma = \omega(\boldsymbol{X})$, 其中 $\boldsymbol{X} \notin \gamma$, 则存在 $\boldsymbol{X}$ 的一个邻域 $\mathcal{O}$ 使得对于所有的 $\boldsymbol{Y} \in \mathcal{O}$, 都有 $\gamma = \omega(\boldsymbol{Y})$. 换句话说就是, 集合

$$\{\boldsymbol{Y} | \omega(\boldsymbol{Y}) = \gamma\} - \gamma$$

是开的. ■

回忆一下, $\mathbb{R}^n$ 中的闭的有界子集称为**紧的**. 我们来看庞加莱-本迪克逊定理的另一个推论, 假设 K 是一个正向不变的集合且是紧的. 如果 $\boldsymbol{X} \in K$, 则 $\omega(\boldsymbol{X})$ 也必定位于 K 中. 从而, K 要么包含一个平衡点要么包含一个极限环.

推论 2 一个正向不变或负向不变的紧集要么包含一个极限环, 要么包含一个平衡点. ■

下面的结果利用了极限环的盘旋性质.

推论 3 令 γ 为一条闭轨, 记 $\mathcal{U}$ 为 γ 包围的开区域. 则 $\mathcal{U}$ 要么包含一个平衡点, 要么包含一个极限环.

证明 记 D 为紧集 $\mathcal{U} \cup \gamma$. 由于 $\mathcal{U}$ 中没有解可以穿过 γ, 因而 D 是不变的. 如果 $\mathcal{U}$ 不包含极限环和平衡点, 则, 对任一 $\boldsymbol{X} \in \mathcal{U}$, 由庞加莱-本迪克逊定理可得,

$$\omega(\boldsymbol{X}) = \alpha(\boldsymbol{X}) = \gamma.$$

如果 $\mathcal{S}$ 是在 γ 上一点 $\boldsymbol{Z}$ 处的局部截面, 则存在序列 $t_n \to \infty, s_n \to -\infty$ 使得 $\phi_{t_n}(\boldsymbol{X}), \phi_{s_n}(\boldsymbol{X}) \in \mathcal{S}$ 且在 $n \to \infty$ 时, $\phi_{t_n}(\boldsymbol{X})$ 和 $\phi_{s_n}(\boldsymbol{X})$ 都趋于 $\boldsymbol{Z}$. 而这与 10.4 节中关于单调序列的命题矛盾. ■

实际上, 这最后一个结果可以大大加强.

推论 4 设 γ 是一条闭轨, 其包围一个开集 U. 则 U 包含一个平衡点.

证明 假设 U 不含平衡点. 先考虑 U 中只有有限个闭轨的情形. 我们可以选取一条闭轨使得它包围的区域面积最小. 于是在这个区域中就没有闭轨或平衡点, 这就与推论 3 矛盾.

现在假设 U 中有无穷多条闭轨. 如果 $\boldsymbol{X}_n$ 都位于闭轨上, 且 $\boldsymbol{X}_n \to \boldsymbol{X}$, 则 $\boldsymbol{X}$ 必定也位于闭轨上. 否则, 由于 U 中没有平衡点, 过 $\boldsymbol{X}$ 的解将盘旋地趋于一个极限环. 由推论 1, 经过附近的 $\boldsymbol{X}_n$ 的解也将趋于这个极限环, 这是不可能的.

记 $v \geqslant 0$ 为 U 中的由闭轨所包围的区域的面积的最大下界. 设 γ_n 为一列闭轨且 $\lim_{n\to\infty} v_n = v$, 其中 v_n 为 γ_n 所围区域的面积. 令 $\boldsymbol{X}_n \in \gamma_n$. 由于 $\gamma \cup U$ 是紧的, 我们不妨假设 $\boldsymbol{X}_n \to \boldsymbol{X} \in U$. 于是如果 U 不含平衡点, 则 $\boldsymbol{X}$ 位于一条闭轨 β 上, 且 β 所围区域的面积为 v. 通常的截面的讨论表明, 当 $n \to \infty$ 时, γ_n 与 β 任

意接近, 从而在 γ_n 和 β 之间的区域的面积 $v_n - v$ 将趋于 0. 上面的讨论证明在 β 内部没有闭轨或平衡点, 而这与推论 3 矛盾. ■

下面的结论以一种很微妙的方式利用了极限环的盘旋性质.

推论 5 令 H 为平面系统的一个首次积分. 如果 H 在任何开集上都不是常数, 则没有极限环.

证明 假设有一个极限环 γ; 令 $c \in \mathbb{R}$ 为 H 在 γ 上的取值 (为常数). 如果 $\boldsymbol{X}(t)$ 是一个盘旋地趋向 γ 的解, 则, 根据 H 的连续性, $H(\boldsymbol{X}(t)) \equiv c$. 在推论 1 中, 我们找到了一个开集使得其上的所有的解都盘旋地趋向 γ; 于是 H 就在一个开集上取常值. ■

最后, 在 9.2 节中发展李雅普诺夫函数的理论时, 就已经隐含了下面的结果.

推论 6 如果 L 是平面系统的一个严格李雅普诺夫函数, 则系统没有极限环. ■

10.7 探索: 振荡的化学反应

在 20 世纪的多数时间, 化学家们相信所有的化学反应都是单调地趋于平衡. 但是, 在 20 世纪 50 年代, 俄罗斯生物化学家 Belousov 发现了一个确定的反应之后, 这个信念就被打破了. 这个反应中包含柠檬酸、溴酸盐离子和硫磺酸, 当它们与一种铈催化剂混合时, 在反应趋于平衡之前, 能振荡相当长的时间. 这种混合物变黄一会儿, 然后变浅, 再变黄, 再变浅, 如此这样可以持续一个多小时. 这种反应, 现在称为 Belousov-Zhabotinsky 反应 (简称 BZ 反应), 它是化学反应历史上的一个主要转折点. 现在, 知道许多反应都会振荡. 有些反应甚至会出现混沌行为.

一个特别简单的这种化学反应是由二氧化氯-碘-丙二酸相互反应给出的. 模拟这个反应的准确的微分方程是极其复杂的. 尽管如此, 存在一个平面非线性系统可以很好地近似其中两个反应物的浓度. 这个系统就是

$$
\begin{aligned}
x' &= a - x - \frac{4xy}{1+x^2} \\
y' &= bx\left(1 - \frac{y}{1+x^2}\right),
\end{aligned}
$$

其中 x, y 分别表示 I^-, ClO_2^- 的浓度, 而 a, b 是正参数.

(1) 探索开始先从数值计算的角度研究这些反应方程. 定性地, 你能找到哪些不同类型的相图?

(2) 找出该系统的所有平衡点.

(3) 在你找到的平衡点处将系统线性化, 并确定每个平衡点的类型.

(4) 在 ab 平面上, 简略地画出你找到的渐近稳定或不稳定平衡点的区域.

(5) 确定系统发生分岔时, a, b 的取值.

(6) 利用系统的零点集和庞加莱-本迪克逊定理找出当存在稳定的极限环时, a, b 的取值. 为什么这些值就对应于振荡的化学反应?

在 [27] 中, 可以看到有关这个反应的更多细节. [47] 则描述了 BZ 反应的非常有趣的历史. [17] 引用了 Belousov 的原始文章.

习　题

1. 对下面的每个系统, 确定所有位于一个 ω 极限集或 α 极限集上的点.

   (a) $r' = r - r^2,\ \theta' = 1$;　(b) $r' = r^3 - 3r^2 + 2r,\ \theta' = 1$;

   (c) $r' = \sin r,\ \theta' = -1$;　(d) $x' = \sin x \sin y,\ y' = -\cos x \cos y$.

2. 考虑三维系统

$$r' = r(1-r)$$
$$\theta' = 1$$
$$z' = -z.$$

   计算沿位于单位圆周 $r = 1, z = 0$ 上的闭轨的庞加莱映射, 并证明这个闭轨是渐近稳定的.

3. 考虑三维系统

$$r' = r(1-r)$$
$$\theta' = 1$$
$$z' = z.$$

   再计算该系统的庞加莱映射. 现在你如何描述闭轨附近的解行为? 简略地作出这个系统的相图.

4. 考虑系统

$$x' = \sin x(-0.1\cos x - \cos y)$$
$$y' = \sin y(\cos x - 0.1\cos y).$$

   证明所有从源点 $(\pi/2, \pi/2)$ 散发出来的解的 ω 极限集都等于 $x = 0, \pi$ 和 $y = 0, \pi$ 围成的正方形.

5. 系统

$$r' = ar + r^3 - r^5$$
$$\theta' = 1$$

   依赖于参数 a. 对于一些有代表性的 a 值确定其相平面, 并描述这个系统的所有分岔.

6. 考虑系统

$$x' = -y - \left(\frac{x^4}{4} - \frac{x^2}{2} + \frac{y^2}{2}\right)(x^3 - x)$$

$$y' = x^3 - x - \left(\frac{x^4}{4} - \frac{x^2}{2} + \frac{y^2}{2}\right) y.$$

(a) 找出所有的平衡点.

(b) 确定这些平衡点的类型.

(c) 证明所有非平衡解的 ω 极限集都由一个鞍点加上一个同宿解或两个同宿解构成.

7. 设 A 为 $\mathbb{R}^2$ 上的一个环形区域, F 为一个平面向量场, 它沿 A 的两条边界曲线都指向 A 的内部. 再假设 A 的每个径向线段都是一个局部截面 (见图 10.12). 证明在 A 中有一个周期解.

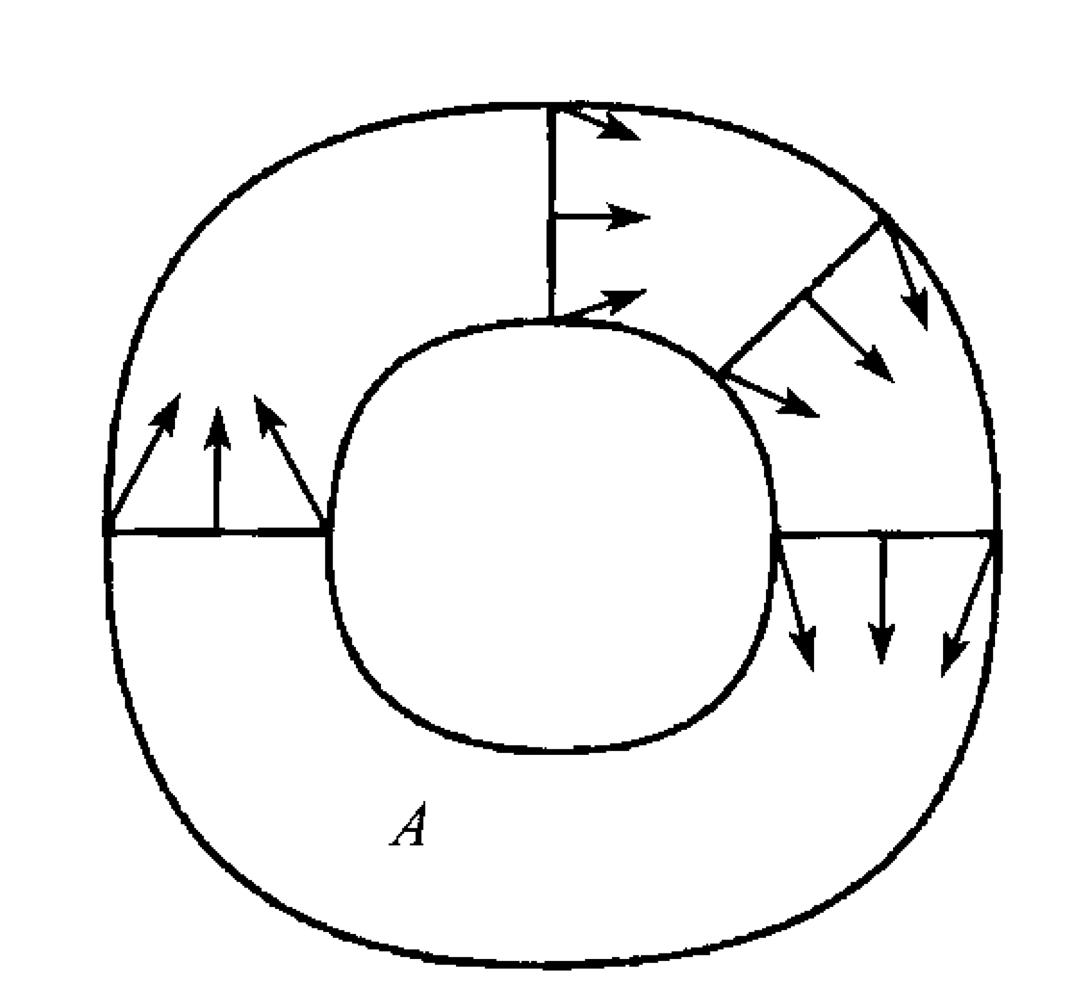


图 10.12　区域 A 是正向不变的

8. 令 F 为一个平面向量场, 仍然考虑上题中的环形区域 A. 同前面一样, 仍然假设 F 没有平衡点, 且沿 A 的两条边界曲线, F 都指向 A 的内部.

(a) 证明在 A 中存在一个闭轨. (注意, 这个假设条件要比上一题的弱.)

(b) 如果在 A 中正好有 7 条闭轨, 证明有一条闭轨使得在它的两边都有轨道盘旋地趋向它.

9. 设 F 是上面的环形区域 A 的一个邻域上的平面向量场. 假设对于 A 的每个边界点 $\boldsymbol{X}$, $F(\boldsymbol{X})$ 都是一个与边界相切的非零向量.

(a) 进一步假设没有平衡点, 且除了边界圆周外没有其他的闭轨, 简略地作出 A 中所有可能的相图. 要包括边界上的解的行进方向相反的情形.

(b) 假设边界解的方向相反, 并且流是保面积的. 证明 A 包含一个平衡点.

10. 证明平面系统的一条闭轨与一个局部截面至多交于一点.

11. 证明一个有界的闭极限集是连通的 (即, 不是两个互不相交的非空闭集的并).

12. 令 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 为一个没有平衡点的平面系统. 假设由 F 生成的流 ϕ_t 保持面积 (即, 如果 U 为任一开集, $\phi_t(U)$ 的面积与 t 无关). 证明每条解曲线都是一个闭集.

13. 设 γ 是平面系统的一条闭轨, λ 是 γ 的周期. 设 $\{\gamma_n\}$ 为一列闭轨. 假设 γ_n 的周期为 λ_n. 如果存在点 $\boldsymbol{X}_n \in \gamma_n$ 使得 $\boldsymbol{X}_n \to \boldsymbol{X} \in \gamma$, 证明 $\lambda_n \to \lambda$. (这一结果对高维系统可能不成立. 然而, 下面的结论成立: 如果 $\lambda_n \to \mu$, 则 μ 是 λ 的整数倍.)

14. 考虑 $\mathbb{R}^2$ 上一个仅有有限个平衡点的系统.

(a) 证明每个极限集要么是一条闭轨, 要么是一些平衡点和一些解 $\phi_t(\boldsymbol{X})$ 的

并, 其中 $\lim_{t\to\infty}\phi_t(\boldsymbol{X})$ 和 $\lim_{t\to-\infty}\phi_t(\boldsymbol{X})$ 都是这些平衡点.

(b) 举例 (画一个图) 说明 $\omega(\boldsymbol{X})$ 中不同解曲线的个数可以是无限的.

15. 令 $\boldsymbol{X}$ 是平面系统的一个**回复点**, 即, 存在序列 $t_n\to\pm\infty$ 使得

$$\phi_{t_n}(\boldsymbol{X})\to\boldsymbol{X}.$$

(a) 证明要么 $\boldsymbol{X}$ 是一个平衡点, 要么 $\boldsymbol{X}$ 位于一条闭轨上.

(b) 举例说明, 对一个非平面系统, 回复点可以既不是平衡点也不在闭轨上.

16. 设 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 和 $\boldsymbol{X}'=\boldsymbol{G}(\boldsymbol{X})$ 为两个平面系统. 假设对所有的 $\boldsymbol{X}\in\mathbb{R}^2$, 都有

$$\boldsymbol{F}(\boldsymbol{X})\cdot\boldsymbol{G}(\boldsymbol{X})=0.$$

如果 $\boldsymbol{F}$ 有一条闭轨, 证明 $\boldsymbol{G}$ 有一个平衡点.

17. 设 γ 是平面系统的一条闭轨, $\mathcal{U}$ 是 γ 内部的有界开区域. 证明 γ 不能同时是 $\mathcal{U}$ 中点的 ω 极限集和 $\mathcal{U}$ 中点的 α 极限集. 利用这个事实以及庞加莱 - 本迪克逊定理证明 $\mathcal{U}$ 包含一个非鞍点的平衡点. (提示: 考虑鞍点的稳定曲线和不稳定曲线上的点的极限集.)

第 11 章　生物学中的应用

在本章，我们利用前几章发展的技巧来考察一些非线性系统，这些系统已经被广泛地用作各种生物系统的数学模型. 在 11.1 节中，我们利用前面关于零点集和线性化的结果来描述有关传染病扩散的几个生物模型. 在 11.2 节，我们研究模拟捕食者/猎物的生态学的最简单的方程. 在 11.3 节，我们用复杂一点的方法来研究一对竞争物种的数量. 我们将仅对这些形式的方程作定性的假设，而不是发展它们的显式表达式. 然后在这些假设条件的基础上，我们得到这类系统的解行为的几何信息.

11.1　传　染　病

像麻疹、疟疾这类传染病的传播可以用一个非线性微分方程系统来模拟. 这种类型的最简单的模型是 SIR 模型. 这里，我们将一个给定的种群分成 3 个不相交的组. 易感种群记为 S，已感种群记为 I，已愈种群记为 R. 通常，这每一个都是时间的函数. 为了简单起见，我们假设种群总量是常数，于是 $(S+I+R)'=0$.

在最基本的情形，我们假设，一个个体被感染而随后又痊愈后就不会再被感染. 像麻疹、腮腺炎、天花以及许多其他疾病都属于这种情形. 我们还假设疾病传播速度正比于易感个体和已感个体的接触次数. 数学上描述这个假设的最简单的方法就是令 $S'=-\beta SI$，其中 $\beta>0$ 为一个常数. 我们最后假设已感个体的治愈速度正比于已被感染的数量. 于是 SIR 模型就是

$$
\begin{aligned}
S' &= -\beta SI \\
I' &= \beta SI - \nu I \\
R' &= \nu I,
\end{aligned}
$$

其中 β,ν 为正参数.

根据约定，我们有 $(S+I+R)'=0$，从而 $S+I+R$ 是一个常数. 如果我们确定了 $S(t),I(t)$，我们就自动地得到了 $R(t)$，这样就简化了系统. 从而，我们只需考虑二维系统

$$
\begin{aligned}
S' &= -\beta SI \\
I' &= \beta SI - \nu I.
\end{aligned}
$$

S 轴 ($I=0$) 上的所有点都是系统的平衡点. 在 $(S,0)$ 处线性化就得到矩阵

$$\begin{pmatrix} 0 & -\beta S \\ 0 & \beta S-\nu \end{pmatrix},$$

从而特征值就是 $0, \beta S-\nu$. 第二个特征值在 $0<S<\nu/\beta$ 时为负, 在 $S>\nu/\beta$ 时为正.

S 零点集为 S 轴和 I 轴. 在 I 轴上, 我们有 $I'=-\nu I$, 从而解沿这条直线简单地趋于原点. I 零点集为 $I=0$ 和竖直线 $S=\nu/\beta$. 于是, 我们就得到零点集的图像, 如图 11.1 所示. 从图中可以看出, 任给初始条件 (S_0, I_0), 如果 $S_0>\nu/\beta, I_0>0$, 则易感个体的数量单调减少, 而已感个体的数量最初先上升, 但在达到一个极大值后将减少到 0.

事实上, 我们可以从分析上证明这件事, 这是因为, 我们可以算出一个沿解曲线取常值的函数. 注意, 这个向量场的斜率只是 S 的函数:

$$\frac{I'}{S'}=\frac{\beta SI-\nu I}{-\beta SI}=-1+\frac{\nu}{\beta S}.$$

于是我们有

$$\frac{\mathrm{d}I}{\mathrm{d}S}=\frac{\mathrm{d}I/\mathrm{d}t}{\mathrm{d}S/\mathrm{d}t}=-1+\frac{\nu}{\beta S},$$

积分后就可得出

$$I=I(S)=-S+\frac{\nu}{\beta}\ln S+常数.$$

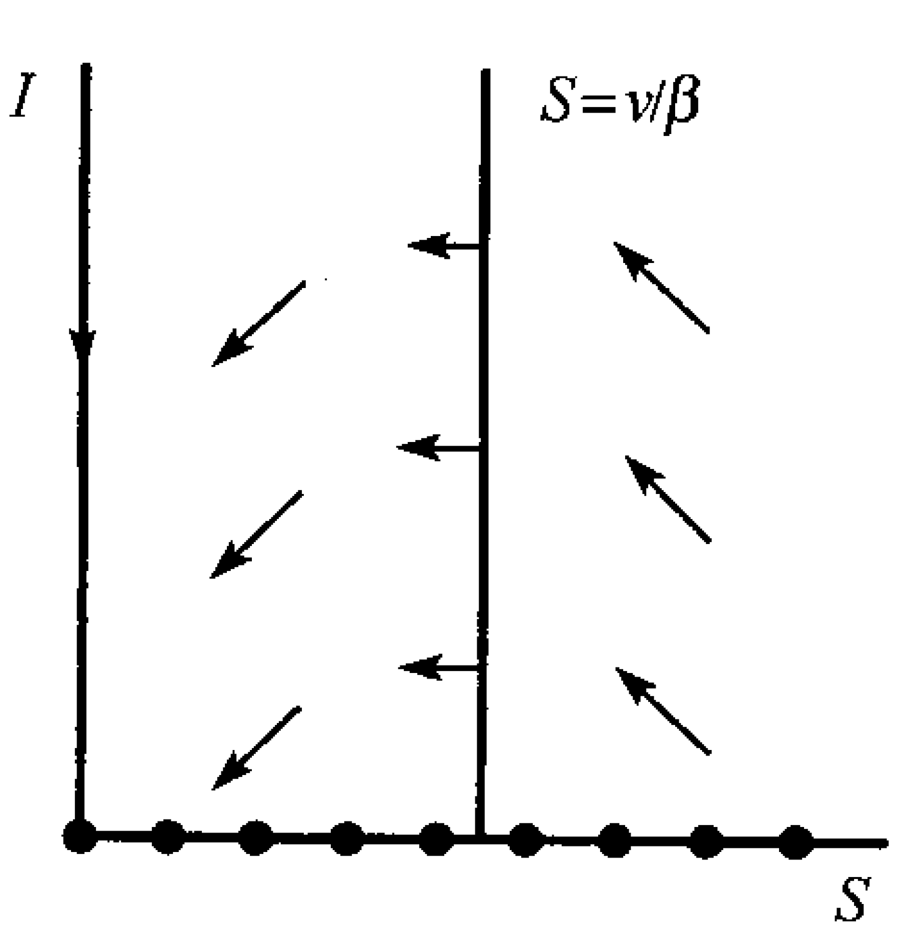


图 11.1 SIR 模型的零点集和方向场

从而函数 $I+S-\nu/\beta \ln S$ 沿解曲线为常数. 由此可得, 对于区间 $\nu/\beta<S<\infty$ 中的每个平衡点, 存在唯一的解曲线将它连接到区间 $0<S<\nu/\beta$ 中的某个平衡点 (见图 11.2).

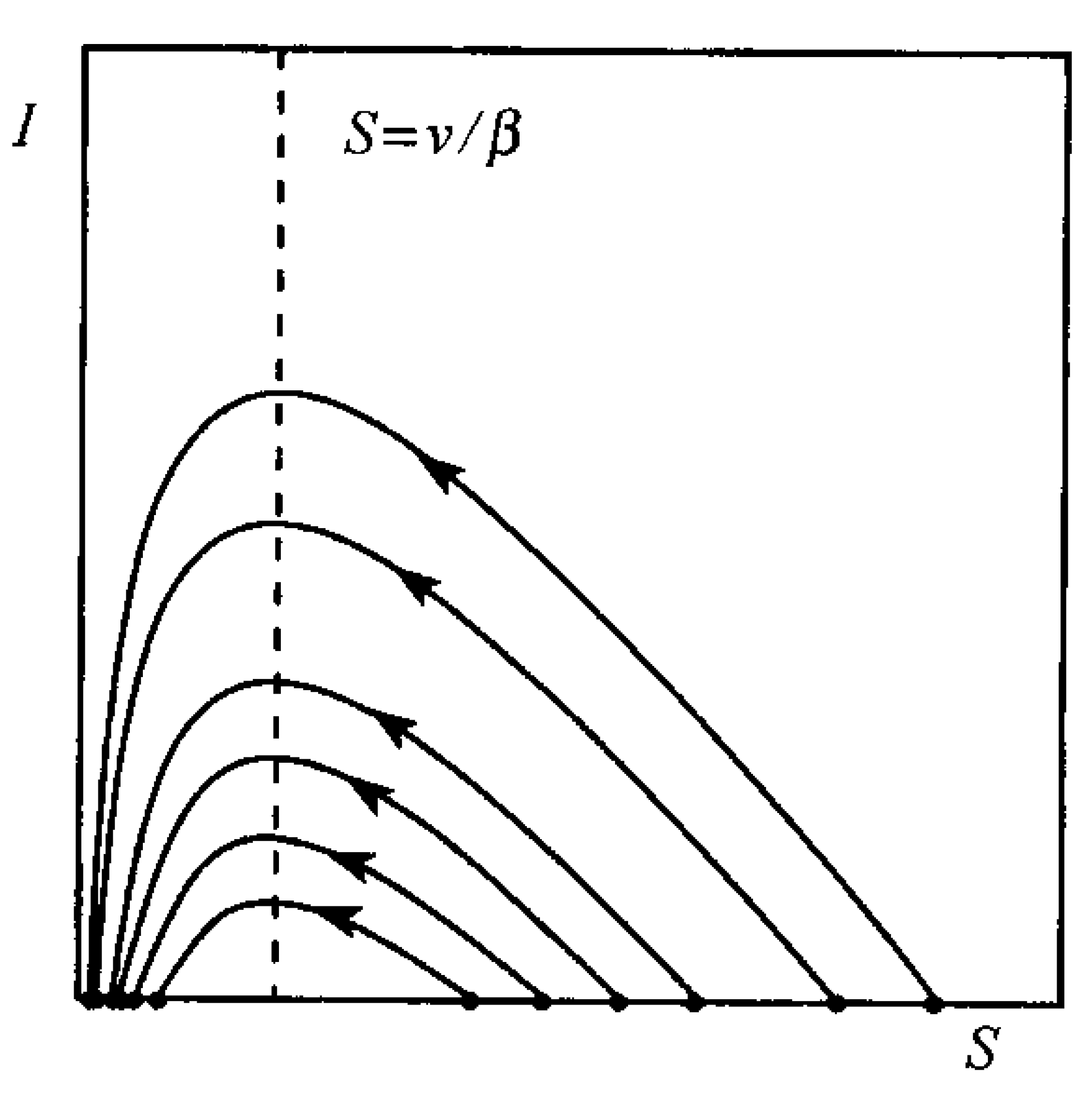


图 11.2 SIR 系统的相图

如果我们假设痊愈的个体可能会失去免疫力而再次感染这种疾病, 这就产生了一个复杂一点的传染病模型. 这种类型的疾病包括疟疾、肺结核等. 我们假设痊愈个体回到 S 类的速度正比于痊愈个体的数量. 这就导致 SIRS 模型 (其中多出来的 S 表示痊愈个体可能再次进入易感群体). 此时系统变成

$$\begin{aligned} S' &= -\beta SI + \mu R \\ I' &= \beta SI - \nu I \\ R' &= \nu I - \mu R. \end{aligned}$$

我们发现种群总量 $S+I+R$ 依然是一个常数, 将它记为 τ. 令 $R=\tau-S-I$, 在这个系统中消去 R 后就得到:

$$\begin{aligned} S' &= -\beta SI + \mu(\tau - S - I) \\ I' &= \beta SI - \nu I. \end{aligned}$$

这里 β,μ,ν,τ 都是正参数.

与 SIR 模型不同, 现在我们至多只有两个平衡点, 一个在 $(\tau,0)$, 另一个在

$$(S^*, I^*) = \left(\frac{\nu}{\beta}, \frac{\mu(\tau-\nu/\beta)}{\nu+\mu}\right).$$

第一个平衡点对应于种群中没有疾病发生. 只有当 $\tau \geqslant \nu/\beta$ 时, 才会出现第二个平衡点. 当 $\tau=\nu/\beta$ 时, 两个平衡点在 $(\tau,0)$ 重合, 从而出现分岔. ν/β 称为这种疾病的阈值.

该系统的线性化系统为

$$\boldsymbol{Y}' = \begin{pmatrix} -\beta I - \mu & -\beta S - \mu \\ \beta I & \beta S - \nu \end{pmatrix} \boldsymbol{Y}.$$

在平衡点 $(\tau,0)$ 处, 两个特征值为 $-\mu, \beta\tau-\nu$, 从而当种群量超过阈值后, 这个平衡点是一个鞍点. 在第二个平衡点处, 直接的计算表明, 矩阵的迹为负, 而它的行列式为正. 从而由第 4 章的结果可知, 两个特征值都有负的实部, 因而这个平衡点是渐近稳定的. 生物学上这意味着只有当种群数量超过阈值时, 这种疾病才可能在这种生物群落中形成. 在后面, 我们将只考虑这种情形.

注意, SIRS 系统有趣的区域只出现在 $I,S\geqslant 0$ 且 $S+I\leqslant\tau$ 上. 记这个三角形区域为 Δ(这是当然!). 注意此时 I 轴不再是不变的, 而在 S 轴上, 解将递增地趋于平衡点 $(\tau,0)$.

命题　区域 Δ 是正向不变的.

证明　我们来检查向量场沿 Δ 边界的方向. 沿下面的边 $I=0$, 向量场与边界相切, 在 $(0,\tau)$ 处, 向量场与竖直方向的边相切. 沿 $S=0$, 我们有 $S'=\mu(\tau-I)>0$, 从而在

$0 < I < \tau$ 时, 向量场指向内部. 沿斜边, 如果 $0 < S \leqslant \nu/\beta$, 我们有 $S' = -\beta SI < 0$ 且 $I' = I(\beta S - \nu) \leqslant 0$, 从而向量场指向内部. 当 $\nu/\beta < S < \tau$ 时, 我们有

$$-1 < \frac{I'}{S'} = -1 + \frac{\nu}{\beta S} \leqslant 0,$$

从而向量场也是指向内部的. 这就完成了证明. ■

与 SIR 模型一样, 此时的 I 零点集由 $I = 0$ 和 $S = \nu/\beta$ 给出. S 零点集为函数

$$I = I(S) = \frac{\mu(\tau - S)}{\beta S + \mu}$$

的图像. 学过微积分的学生都可以算出, 在 $0 \leqslant S < \tau$ 时, $I'(S) < 0$, $I''(S) > 0$. 从而这个零点集是一个经过 $(\tau, 0)$ 和 $(0, \tau)$, 单调下降且下凹的函数的图像, 如图 11.3 所示. 注意, 从这个相图中可以看到, 所有的解都趋于平衡点 (S^*, I^*); 而感染个体与易感个体的比例趋于一个"稳定的状态". 然而, 为了证明这点, 对于一组给定的参数 β, μ, ν, τ, 还需要说明不会出现包围这个平衡点的闭轨.

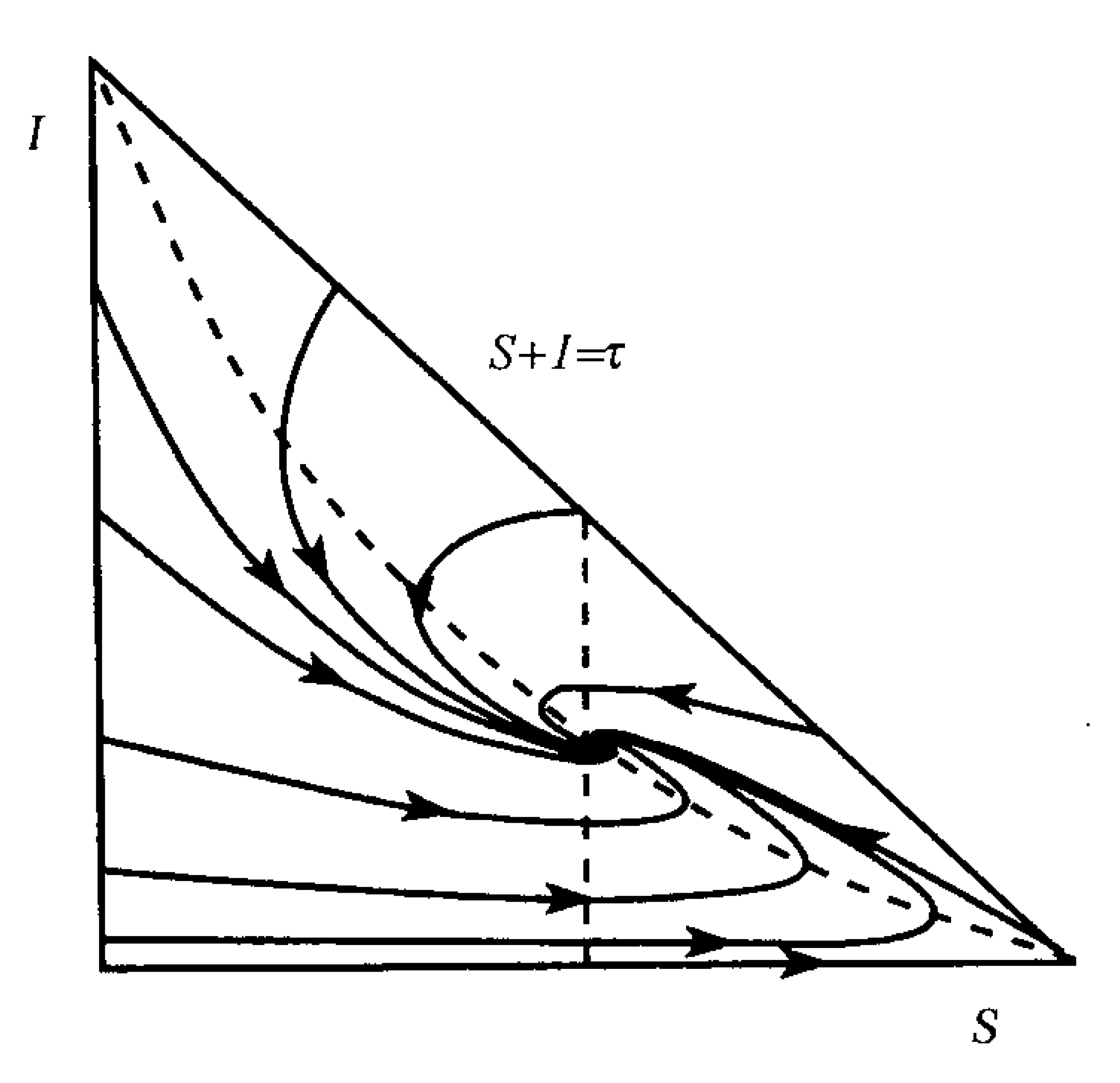


图 11.3 SIRS 系统在 Δ 中的零点集和相图. 这里 $\beta = \nu = \mu = 1, \tau = 2$

11.2 捕食者/猎物系统

下面我们考虑一对种群, 其中一个由捕食者构成, 它们的数量记为 y, 另外一种是它们的猎物, 数量为 x. 我们假设这种猎物就是捕食者的全部食物. 我们还假设, 在没有捕食者时, 猎物数量的增长速度正比于当前的猎物数量. 即, 在 $y = 0$ 时, 与第 1 章一样, 我们有 $x' = ax$, 其中 $a > 0$. 从而在这种情形 $x(t) = x_0 \exp(at)$. 当捕食者出现后, 我们假设猎物数量的减少速度正比于捕食者/猎物相遇的次数. 与

上节一样, 这种描述的一个简单模型就是 bxy, 其中 $b>0$. 从而猎物数量的微分方程就是 $x'=ax-bxy$.

对于捕食者的数量, 我们作的假设或多或少是相反的. 在没有猎物时, 捕食者数量的减少速度正比于当前的数量. 也就是, 在 $x=0$ 时, 我们有 $y'=-cy$, 其中 $c>0$, 于是 $y(t)=y_0\exp(-ct)$. 在这种情形, 捕食者种群将灭绝. 当环境中有猎物时, 我们假设捕食者数量的增长速度正比于捕食者/猎物相遇的次数, 也就是 dxy. 此时, 我们对过度拥挤时不作任何假设. 于是, 我们就得到简化的捕食者/猎物系统 (也称为 Volterra-Lotka 系统)

$$\begin{aligned} x' &= ax-bxy = x(a-by) \\ y' &= -cy+dxy = y(-c+dx), \end{aligned}$$

其中参数 a,b,c,d 都假设是正的. 由于现在考虑的是种群数量, 我们只关心 $x,y\geqslant 0$.

与通常一样, 第一步工作是确定平衡点的位置. 此时的平衡点出现在原点和 $(x,y)=(c/d,a/b)$. 线性化系统为

$$\boldsymbol{X}'=\begin{pmatrix} a-by & -bx \\ dy & -c+dx \end{pmatrix}\boldsymbol{X}.$$

于是, 当 $x=y=0$ 时, 我们得到一个特征值为 a 和 $-c$ 的鞍点. 而且我们知道它的稳定曲线和不稳定曲线分别就是 y 轴和 x 轴.

在另一个平衡点 $(c/d,a/b)$ 处, 特征值是纯虚数 $\pm\mathrm{i}\sqrt{ac}$, 因而此时关于平衡点的稳定性, 我们不能得出任何结论.

下面我们再简略地作出该系统的零点集. 其 x 零点集由直线 $x=0$ 和 $y=a/b$ 给出, 而其 y 零点集为 $y=0$ 和 $x=c/d$. 非零的零点集中的两条直线将区域 $x,y>0$ 分成 4 个基本区域, 其中的向量场的指向如图 11.4 所示. 于是, 解围着平衡点反时针方向缠绕.

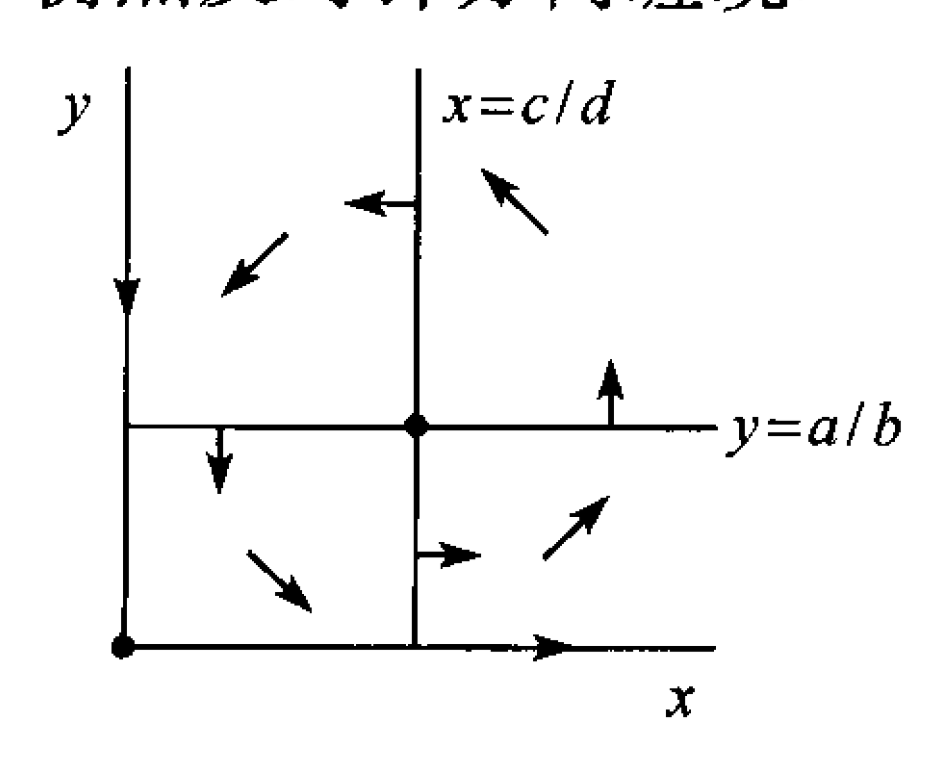


图 11.4 捕食者 / 猎物系统的零点集和方向场

从这点, 我们还是不能确定解的准确行为: 它们可能盘旋地进入平衡点, 可能盘旋地趋向一个极限环, 可能顺着坐标轴盘旋地趋向"无穷", 也有可能位于闭轨上. 为了作出判定, 我们来寻找一个李雅普诺夫函数 L. 利用**分离变量**的技巧, 我们来寻找一个如下形式的函数

$$L(x,y)=F(x)+G(y).$$

回忆一下, $\dot{L}$ 代表 L 沿解对时间的导数. 我们来作计算

$$\dot{L}(x,y)=\frac{\mathrm{d}}{\mathrm{d}t}L(x(t),y(t))=\frac{\mathrm{d}F}{\mathrm{d}x}x'+\frac{\mathrm{d}G}{\mathrm{d}y}y'.$$

于是

$$\dot{L}(x,y) = x\frac{\mathrm{d}F}{\mathrm{d}x}(a-by) + y\frac{\mathrm{d}G}{\mathrm{d}y}(-c+dx).$$

如果

$$\frac{x\,\mathrm{d}F/\mathrm{d}x}{dx-c} = \frac{y\,\mathrm{d}G/\mathrm{d}y}{by-a},$$

则我们就可得出 $\dot{L} \equiv 0$. 由于 x, y 为独立的变量, 上式成立当且仅当

$$\frac{x\,\mathrm{d}F/\mathrm{d}x}{dx-c} = \frac{y\,\mathrm{d}G/\mathrm{d}y}{by-a} = \text{常数}.$$

令这个常数为 1, 我们得到

$$\frac{\mathrm{d}F}{\mathrm{d}x} = d - \frac{c}{x},$$
$$\frac{\mathrm{d}G}{\mathrm{d}y} = b - \frac{a}{y}.$$

积分之后, 我们就找到

$$F(x) = dx - c\ln x,$$
$$G(y) = by - a\ln y.$$

从而当 $x, y > 0$ 时, 函数

$$L(x,y) = dx - c\ln x + by - a\ln y$$

沿该系统的解曲线为常数.

通过考察 $\partial L/\partial x, \partial L/\partial y$ 的符号, 容易看出平衡点 $\boldsymbol{Z} = (c/d, a/b)$ 是 L 的一个严格极小值. 由此可得 L(或, 更严格地说, $L - L(\boldsymbol{Z})$) 就是该系统的一个李雅普诺夫函数. 从而, $\boldsymbol{Z}$ 是一个稳定的平衡点.

然后我们注意到该系统没有极限环; 这是因为 L 在任何开集上都不是常数, 从而由 10.6 节中的推论 5 立得. 现在我们证明下面的定理.

定理 捕食者/猎物系统的每个解都是闭轨 (平衡点 $\boldsymbol{Z}$ 和坐标轴除外).

证明 考虑过 $\boldsymbol{W} \neq \boldsymbol{Z}$ 的解, 其中 $\boldsymbol{W}$ 不在 x 轴或 y 轴上. 这个解围着 $\boldsymbol{Z}$ 点盘旋, 且无穷次穿过每个零点集. 从而有一个双边无穷序列 $\cdots < t_{-1} < t_0 < t_1 < \cdots$, 使得 $\phi_{t_n}(\boldsymbol{W})$ 位于直线 $x = c/d$ 上, 且当 $n \to \pm\infty$ 时, $t_n \to \pm\infty$. 如果 $\boldsymbol{W}$ 不在闭轨上, 像上一章一样的讨论可以得出, 点列 $\phi_{t_n}(\boldsymbol{W})$ 沿直线 $x = c/d$ 是单调的. 由于没有极限环, 因而要么在 $n \to \infty$ 时, $\phi_{t_n}(\boldsymbol{W}) \to \boldsymbol{Z}$, 要么在 $n \to -\infty$ 时, $\phi_{t_n}(\boldsymbol{W}) \to \boldsymbol{Z}$. 由于 L 沿过 $\boldsymbol{W}$ 的解为常数, 这就蕴涵 $L(\boldsymbol{W}) = L(\boldsymbol{Z})$. 但这与 $L(\boldsymbol{Z})$ 的严格极小性矛盾. 这就完成了证明. ■

捕食者/猎物系统的相图如图 11.5 所示. 我们得出, 对于任意给定的初始条件 $(x(0), y(0))$, $x(0) \neq 0, y(0) \neq 0$, 且不是 $\boldsymbol{Z}$, 捕食者和猎物的数量是循环振荡的. 无论捕食者和猎物系统的数量是多少, 都没有种群会灭绝, 也没有种群的数量会无限制地增长.

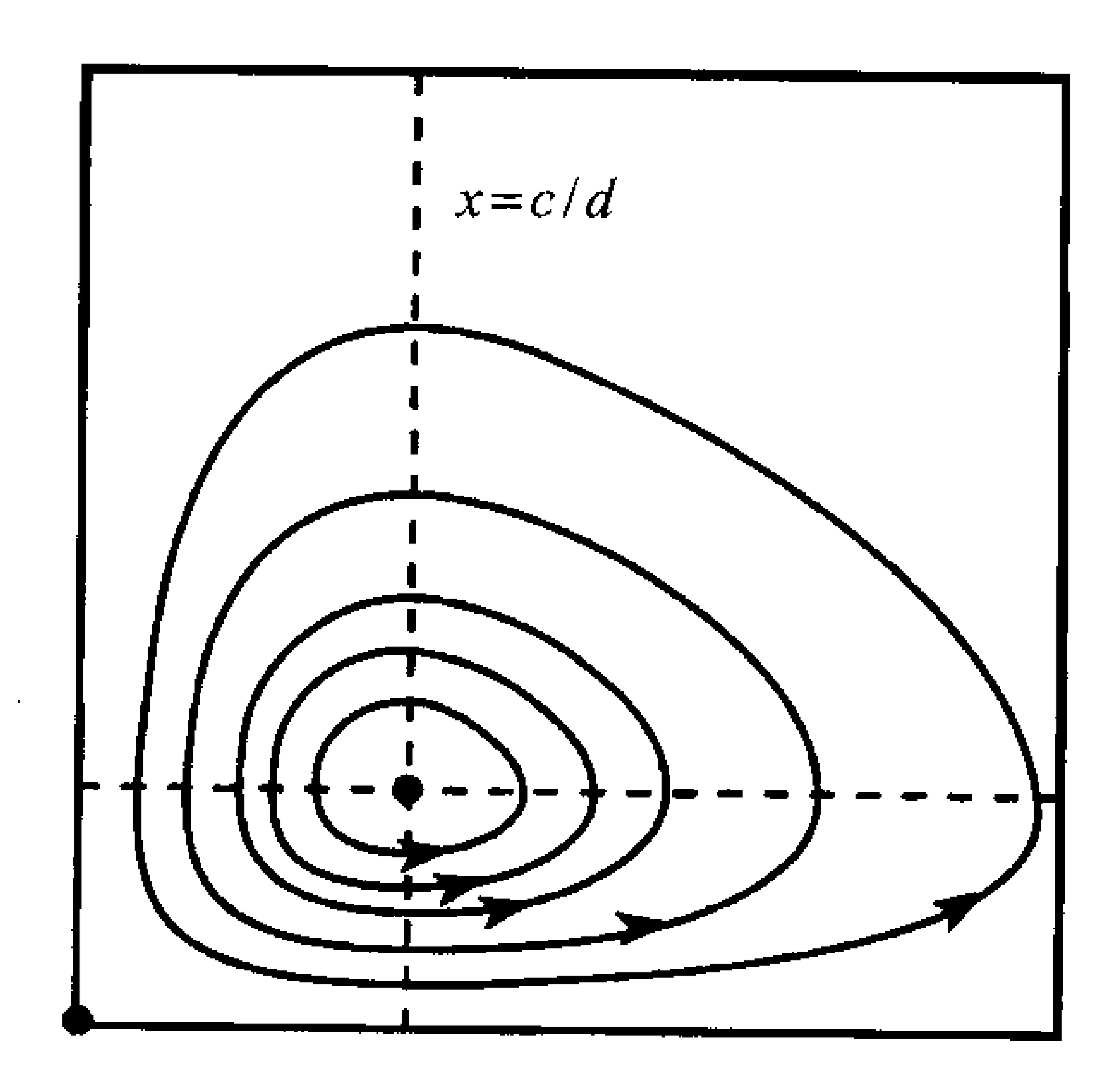


图 11.5 捕食者/猎物系统的零点集和相图

现在我们在猎物方程中引入过度拥挤. 与第 1 章的合理模型一样, 在没有捕食者时, 猎物方程可以写成

$$x' = ax - \lambda x^2.$$

我们还假设在 $x = 0$ 时, 捕食者数量服从一个类似的方程

$$y' = -cy - \mu y^2.$$

将上面的假设考虑进来后就得到**种群具有极限增长时的捕食者/猎物方程**:

$$\begin{aligned} x' &= x(a - by - \lambda x) \\ y' &= y(-c + dx - \mu y). \end{aligned}$$

像以前一样, 参数 a, b, c, d, λ, μ 都是正的. 当 $y = 0$ 时, 我们得到合理方程 $x' = x(a - \lambda x)$, 从而得到位于原点和 $(a/\lambda, 0)$ 的平衡点. 在第 1 章我们已经看到, x 轴上的所有非零解都趋于 a/λ.

当 $x = 0$ 时, 关于 y 的方程为 $y' = -cy - \mu y^2$. 由于当 $y > 0$ 时, $y' > 0$, 因而这个轴上的所有解都趋于原点. 于是, 我们将注意力限制在第一象限 Q, 也就是 $x, y > 0$.

零点集由 x 轴, y 轴以及直线

$$\begin{aligned} L:&\ a - by - \lambda x = 0, \\ M:&\ -c + dx - \mu y = 0 \end{aligned}$$

构成. 沿直线 L 和 M, 我们分别有 $x'=0$ 和 $y'=0$. 根据这两条直线在 Q 中相交与否, 我们将讨论分成两种情况.

我们首先考虑这两条直线在 Q 中不相交的情形. 此时, 零点集的图像如图 11.6 所示. 所有在 M 右方的解都向左上方前进直到碰到 M; 在两条直线 L, M 之间的解向左下方前进. 于是, 它们要么碰到 L, 要么直接趋于平衡点 $(a/\lambda, 0)$. 如果解穿过 L, 它们将向右下方前进, 但它们不会再次穿过 L. 从而它们也趋于 $(a/\lambda, 0)$. 于是 Q 中所有的解都趋于这个平衡点. 我们就得出, 在这种情形, 捕食者的种群将逐渐灭绝, 而猎物的数量则趋于其极限值 a/λ.

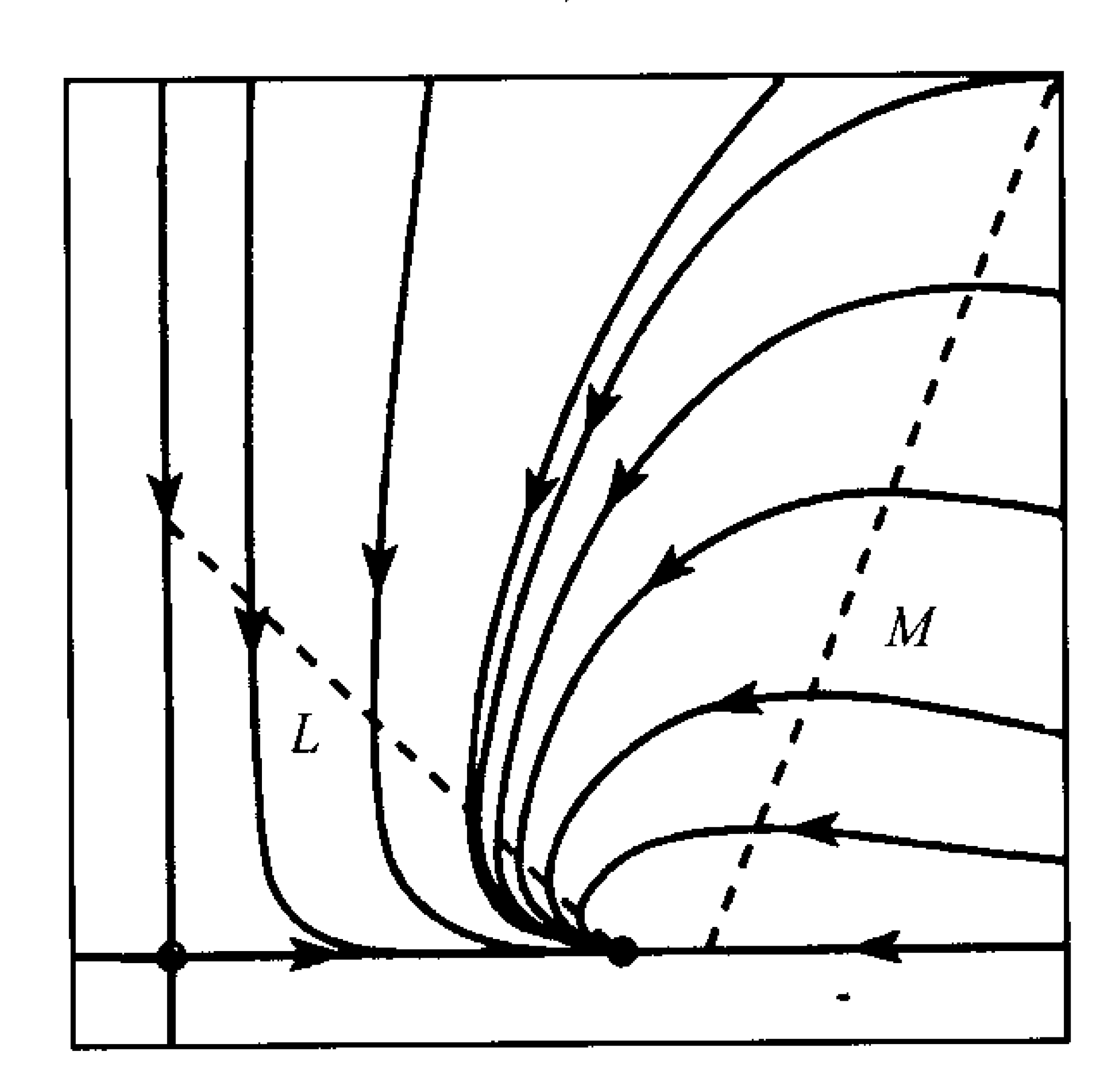


图 11.6 具有极限增长时的捕食者/猎物系统的零点集和相图:
零点集在 Q 中不相交的情形

我们可对零点集附近的解行为作如下解释. 由于 x', y' 不可能同时为正, 因而捕食者和猎物不可能同时增长. 如果猎物数量超过其极限值, 它就必须减少. 一段时间后, 猎物的减少将会使得捕食者数量开始减少 (当解穿过 M). 过了这个点后, 猎物的数量就再也不能超过 a/λ, 从而捕食者的数量继续减少. 如果解穿过 L, 猎物的数量再次增长 (但不超过 a/λ), 于是捕食者继续死亡. 在极限情形, 捕食者消失, 而猎物的数量稳定在 a/λ.

现在假设 L 和 M 在 Q 中有一个交点 $\boldsymbol{Z}=(x_0, y_0)$; 当然 $\boldsymbol{Z}$ 是一个平衡点. 向量场在 $\boldsymbol{Z}$ 处的线性化为

$$\boldsymbol{X}' = \begin{pmatrix} -\lambda x_0 & -bx_0 \\ dy_0 & -\mu y_0 \end{pmatrix} \boldsymbol{X}.$$

其中的矩阵的迹为 $-\lambda x_0 - \mu y_0 < 0$, 行列式为 $(bd+\lambda\mu)x_0 y_0 > 0$. 从第 4 章的迹–行列式平面可知, $\boldsymbol{Z}$ 的两个特征值要么都是负的, 要么都是具有负实部的复数. 从而 $\boldsymbol{Z}$ 是渐近稳定的.

注意, 除了位于 $\boldsymbol{Z}$ 和 $(0,0)$ 处的平衡点外, 在 $(a/\lambda,0)$ 处还有一个平衡点. 利用线性化可以证明这个平衡点是一个鞍点; 它的稳定曲线位于 x 轴上. 见图 11.7.

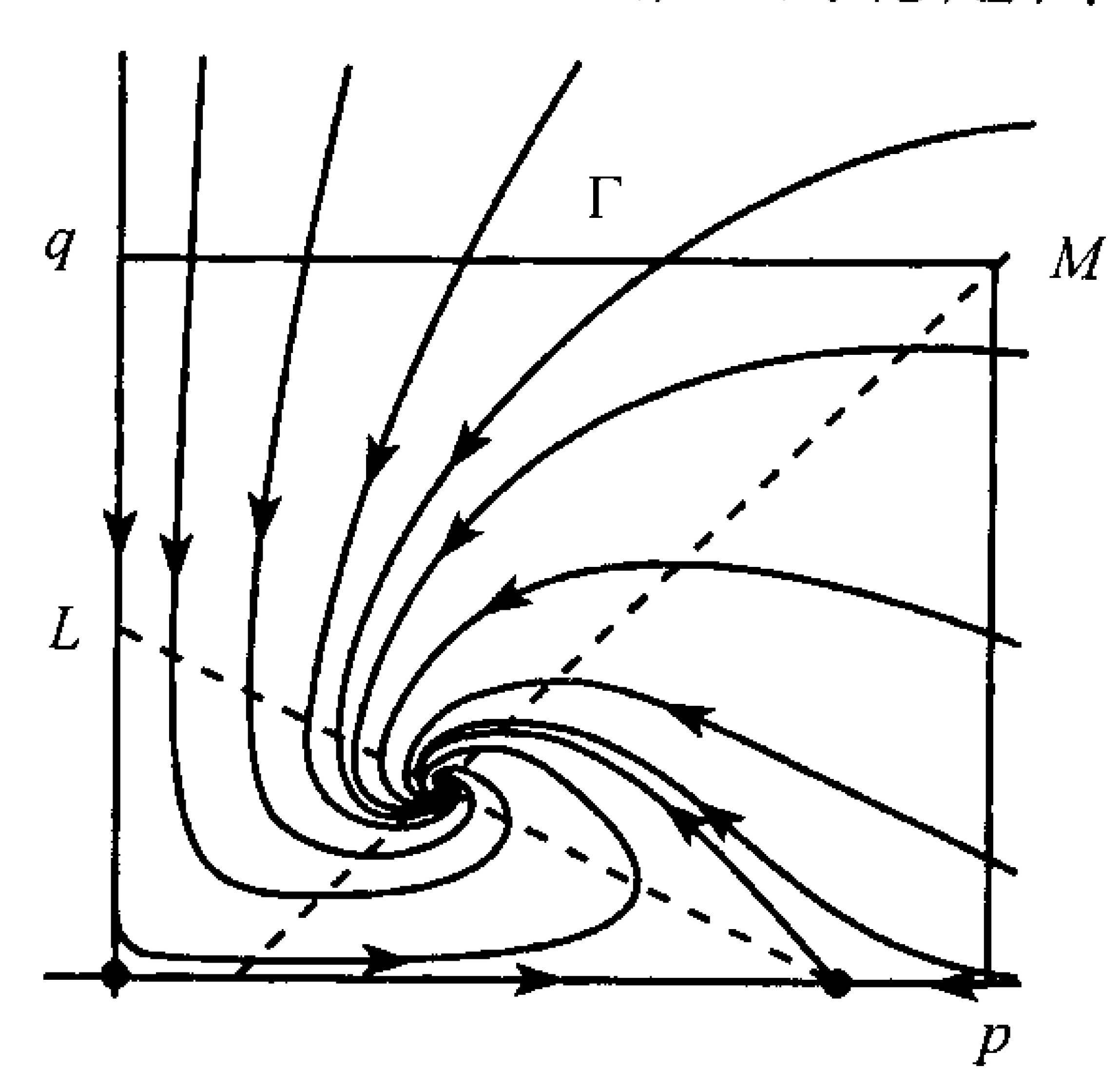


图 11.7 具有极限增长时的捕食者/猎物系统的零点集和相图: 零点集在 $\mathcal{Q}$ 中相交的情形

确定 $\boldsymbol{Z}$ 的吸引盆并不容易, 而且我们也不知道是否有极限环. 不过我们还是能得到某些信息. 直线 L 与 x 轴交于 $(a/\lambda,0)$, 而与 y 轴交于 $(0,a/b)$. 令 Γ 为一个矩形, 它的四个顶点分别位于 $(0,0),(p,0),(0,q),(p,q)$, 其中 $p>a/\lambda, q>a/b$, 而且点 (p,q) 位于直线 M 上. 从 Γ 的边界点出发的每个解要么进入 Γ, 要么是边界的一部分. 从而 Γ 是正向不变的. $\mathcal{Q}$ 中的每个点都包含在一个这样的矩形中.

根据庞加莱-本迪克逊定理, Γ 中任一点 $(x,y)(x,y>0)$ 的 ω 极限集必定是一个极限环, 或者必定包含三个平衡点 $(0,0), \boldsymbol{Z}, (a/\lambda,0)$ 中的一个. 注意到平衡点 $(0,0)$ 和 $(a/\lambda,0)$ 都是鞍点, 它们的稳定曲线分别位于 x 轴和 y 轴上, 因而我们可以排除这两个平衡点. 从而 $\omega(x,y)$ 要么是 $\boldsymbol{Z}$, 要么是 Γ 中的一个极限环. 由庞加莱-本迪克逊定理的推论 4, 任何极限环必定包围 $\boldsymbol{Z}$.

我们进一步观察到任何这样的矩形 Γ 都包含所有的极限环, 这是因为一个极限环 (像任何解一样) 一定进入 Γ, 而且 Γ 是正向不变的. 如上固定 (p,q), 于是, 对任何的初始条件 $(x(0),y(0))$, 存在 $t_0>0$, 使得当 $t\geqslant t_0$ 时, $x(t)<p, y(t)<q$. 我们得出, 最后, 解要么趋于 $\boldsymbol{Z}$, 要么盘旋地趋向一个极限环.

从实用的观点看, 某个时间后, 趋于 $\boldsymbol{Z}$ 的解和 $\boldsymbol{Z}$ 就无法区分. 同样, 趋于极限环 γ 的解与 γ 充分接近时就与 γ 相同. 我们得出, 服从这些微分方程的捕食者和猎物的数量要么稳定到一个常数, 要么稳定到一个周期变化的量. 而且, 无论初始的物种数量是多少, 最后的物种数量都不会超过确定的上界.

11.3 竞争物种

现在我们考虑两个物种竞争同一种食物. 我们将不分析特定的方程, 而是采

取一种不同的方法：考虑一大类的方程, 而只假设很少的定性特征. 之所以采用这种方法是由于, 特定的方程可能很难分析, 而且用这种方法可以得到相当的一般性, 而丢失的东西又很少.

记 x, y 为两个物种的数量. 描述两个物种数量增长的方程可以写成如下的形式：

$$\begin{aligned} x' &= M(x,y)x, \\ y' &= N(x,y)y, \end{aligned}$$

其中增长率 M, N 同时是两个变量的函数. 我们照常假设 x, y 是非负的. 从而, x 零点集由 $x=0$ 和 $M(x,y)=0$ 给出, y 零点集由 $y=0$ 和 $N(x,y)=0$ 给出. 我们对 M 和 N 作下面的一些假设：

(1) 由于两个物种竞争同一资源, 当其中一个物种的数量增加时, 另一个物种的增长率就会下降. 于是,

$$\frac{\partial M}{\partial y} < 0 \quad 且 \quad \frac{\partial N}{\partial x} < 0.$$

(2) 如果其中一个物种量非常大, 则两个物种的数量都减少. 从而存在 $K > 0$ 使得当 $x \geqslant K$ 或 $y \geqslant K$ 时,

$$M(x,y) < 0 \quad 且 \quad N(x,y) < 0.$$

(3) 当其中一个物种不出现时, 另一个物种在某个数量之前有正的增长率, 而之后的增长率为负. 从而, 存在常数 $a, b > 0$ 使得

$$\begin{aligned} &当\ x < a\ 时,\ M(x,0) > 0, \quad 且 \quad 当\ x > a\ 时,\ M(x,0) < 0, \\ &当\ y < b\ 时,\ N(0,y) > 0, \quad 且 \quad 当\ y > b\ 时,\ N(0,y) < 0. \end{aligned}$$

由条件 (1) 和条件 (3), 在 $0 \leqslant x \leqslant a$ 时, 每条竖直线 $\{x\} \times \mathbb{R}$ 与集合 $\mu = M^{-1}(0)$ 正好相交一次, 而在 $x > a$ 时, 它们根本不相交. 由条件 (1) 和隐函数定理, 存在非负函数 $f : [0,a] \to \mathbb{R}, f^{-1}(0) = a$ 使得 μ 为 f 的图像. 在曲线 μ 的下方, M 为正, 在 μ 的上方, M 为负. 用同样的方式可得, 集合 $\nu = N^{-1}(0)$ 是一条形如

$$\{(x,y) | x = g(y)\}$$

的光滑曲线, 其中 $g : [0,b] \to \mathbb{R}$ 是一个非负函数, 且 $g^{-1}(0) = b$. 在 ν 的左边, 函数 N 为正, 而在 ν 的右边, N 为负.

先假设 μ, ν 不相交, 且 μ 在 ν 的下方. 此时系统的相图可以通过它的零点集立刻确定. 平衡点为 $(0,0), (a,0), (0,b)$. 位于原点的平衡点是一个源点, 而 $(a,0)$ 是一个鞍点 (假设 $(\partial M/\partial x)(a,0) < 0$). 位于 $(0,b)$ 的平衡点是一个汇点 [再次假设

$(\partial N/\partial y)(0,b)<0$]. 除了 x 轴上的解外, 换句话说, 当 $y_0>0$ 时, 所有从 (x_0,y_0) 出发的解都趋于渐近稳定的平衡点 $(0,b)$(见图 11.8). 在 μ 位于 ν 上方的情形, 情况正好反过来：满足 $x_0>0$ 的所有解都趋于汇点, 不过此时汇点出现在 $(a,0)$.

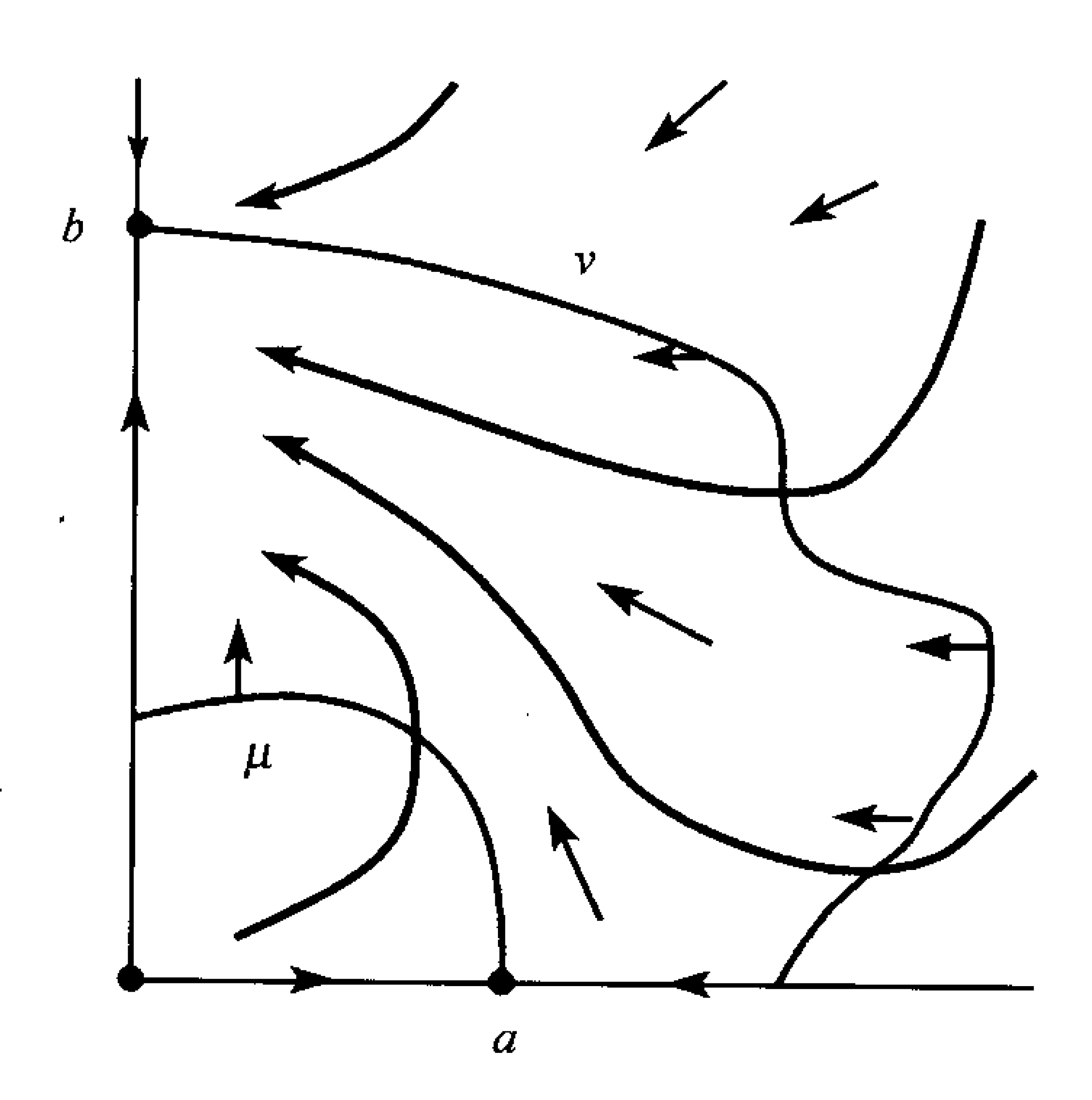


图 11.8 当 μ 和 ν 不相交时的相图

现在假设 μ 和 ν 相交. 我们假设, $\mu\cap\nu$ 是一个有限集, 而且在每个交点处, μ 和 ν 都是**横截地交叉**, 即, 在这些交点处, 它们有不同的切线. 这个假设可以去掉; 我们作这个假设只是为了简化确定这个流的过程.

零点集 μ,ν 以及两个坐标轴在第一象限中围成了有限个连通的开集：这些就是该系统的基本区域, 在这些区域中, $x'\neq 0, y'\neq 0$. 它们可以分成四种类型：

$$\text{A}:\quad x'>0,\ y'>0,\qquad \text{B}:\quad x'<0,\ y'>0,$$

$$\text{C}:\quad x'<0,\ y'<0,\qquad \text{D}:\quad x'>0,\ y'<0.$$

换句话说, 在这四种区域中, 向量场分别指向东北、西北、西南、东南. 一些这种区域如图 11.9 所示. 基本区域 $\mathcal{R}$ 的边界 $\partial\mathcal{R}$ 由以下几类点构成：$\mu\cap\nu$ 中的点, 称为**顶点**; 位于 μ 或 ν 上的点, 但不同时在它们上面, 而且也不在坐标轴上, 称为**正常边界点**; 以及坐标轴上的点.

一个顶点就是一个平衡点; 其他的平衡点位于坐标轴上的 $(0,0),(a,0),(0,b)$. 在一正常边界点 $Z\in\partial\mathcal{R}$, 向量场要么是竖直的 (如果 $Z\in\mu$), 要么是水平的 (如果 $Z\in\nu$). 因为 μ 没有竖直切线, 而 ν 没有水平切线, 所以向量场要么指向 $\mathcal{R}$ 的内部, 要么指向 $\mathcal{R}$ 的外部. 我们将 Z 相应地称为 $\partial\mathcal{R}$ 的**向内的点**或**向外的点**. 注意, 在图 11.9 中, 在一个基本区域的所有正常边界点, 向量场要么全部指向内部, 要么全部指向外部. 这不是偶然的, 因为我们有：

命题 设 $\mathcal{R}$ 是竞争物种模型的一个基本区域. 则 $\mathcal{R}$ 的正常边界点要么全是向内的, 要么全是向外的.

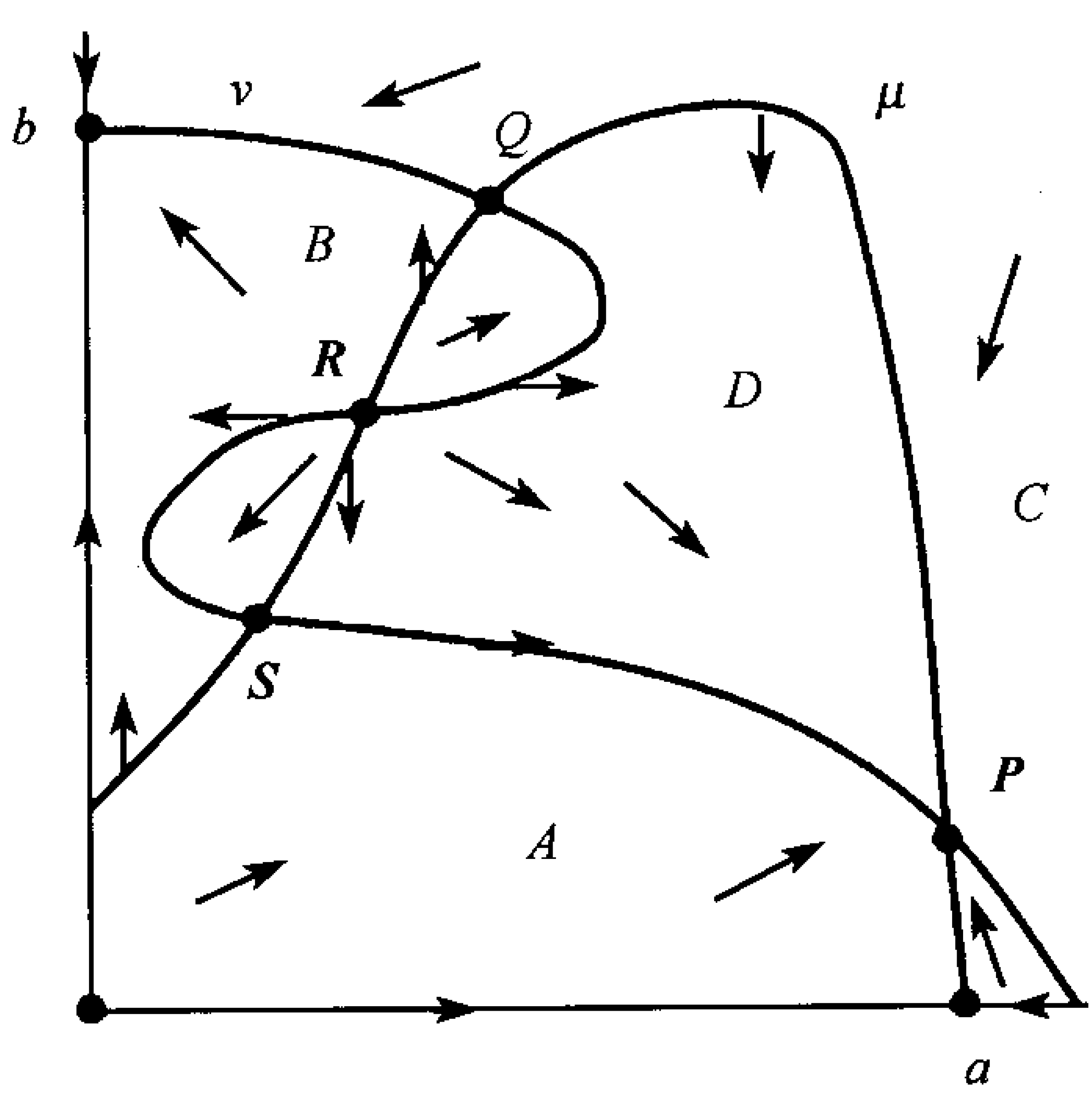


图 11.9 当 μ 和 ν 相交时的基本区域

证明 曲线 μ 和 ν 在一个顶点 $\boldsymbol{P}$ 处相交的方式只有两种可能. 当 y 沿曲线 ν 增加时, ν 要么是从 μ 的下方穿到上方, 要么是从 μ 的上方穿到下方. 这两种模式分别如图 11.10 中的 (a), (b) 所示. 因为我们已经假设两条曲线横截相交, 从而没有其他的可能.

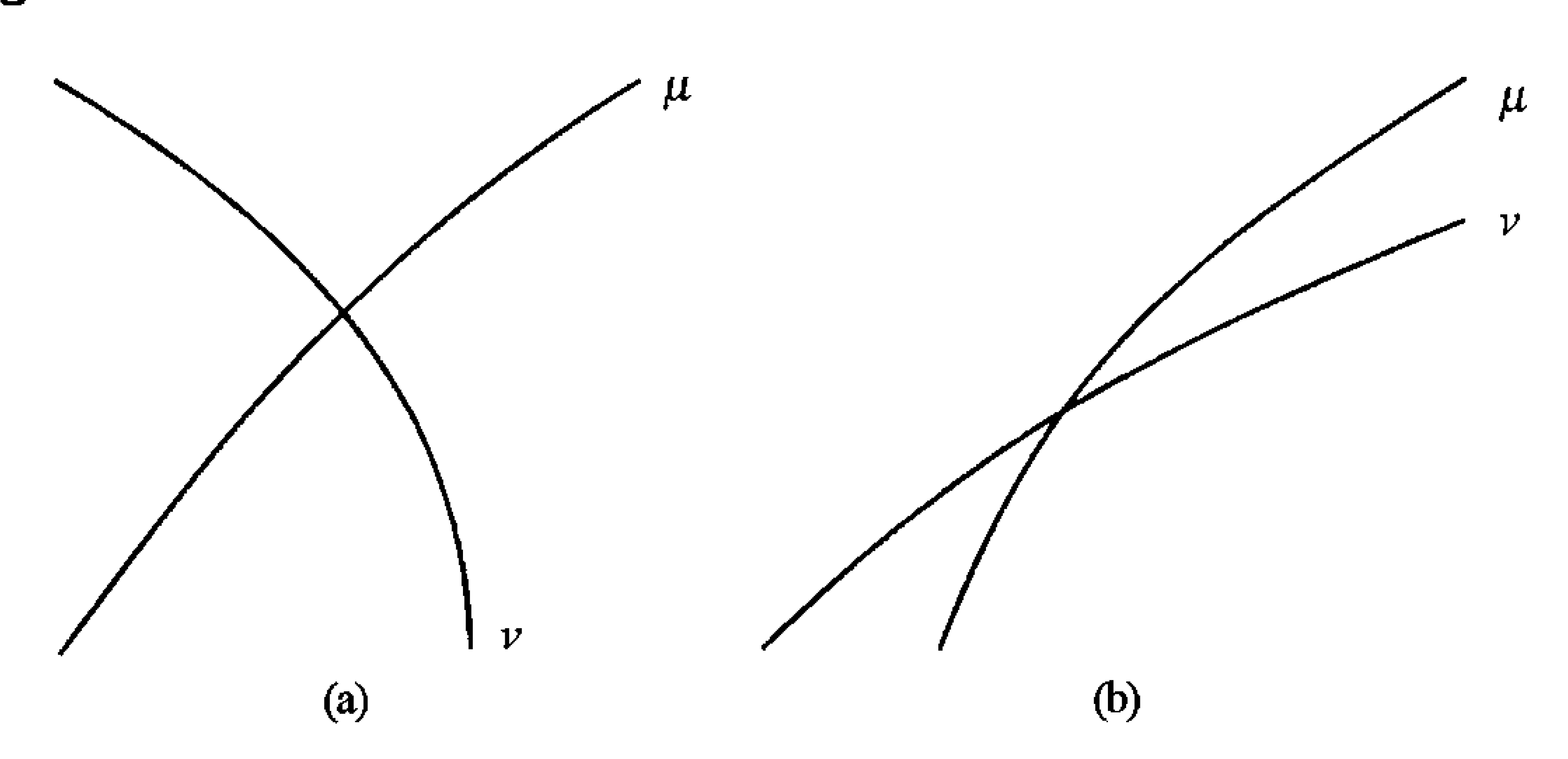


图 11.10 在 (a) 中, 当 y 增加时, ν 从 μ 的下方穿到上方. 在 (b) 中情况正好倒过来

由于在 μ 的下方 $x' > 0$, 在 μ 的上方 $x' < 0$; 且在 ν 的左边 $y' > 0$, 在 ν 的右边 $y' < 0$, 因而在这两种情形, 向量场的结构如图 11.11 所示.

在每种情形, 在关于 $\boldsymbol{P}$ 相对的两个基本区域中, 向量场指向内部, 而在另外两个相对的基本区域中, 向量场指向外部.

如果我们沿着 μ 或 ν 移动到这个基本区域的下一个顶点, 我们看到, 在相邻的两个基本区域中, 向量场保持指向内部或指向外部的结构. 于是, 在每个基本区

域的所有正常边界点处，向量场要么都指向外部，要么都指向内部，这正是我们要证的. ■

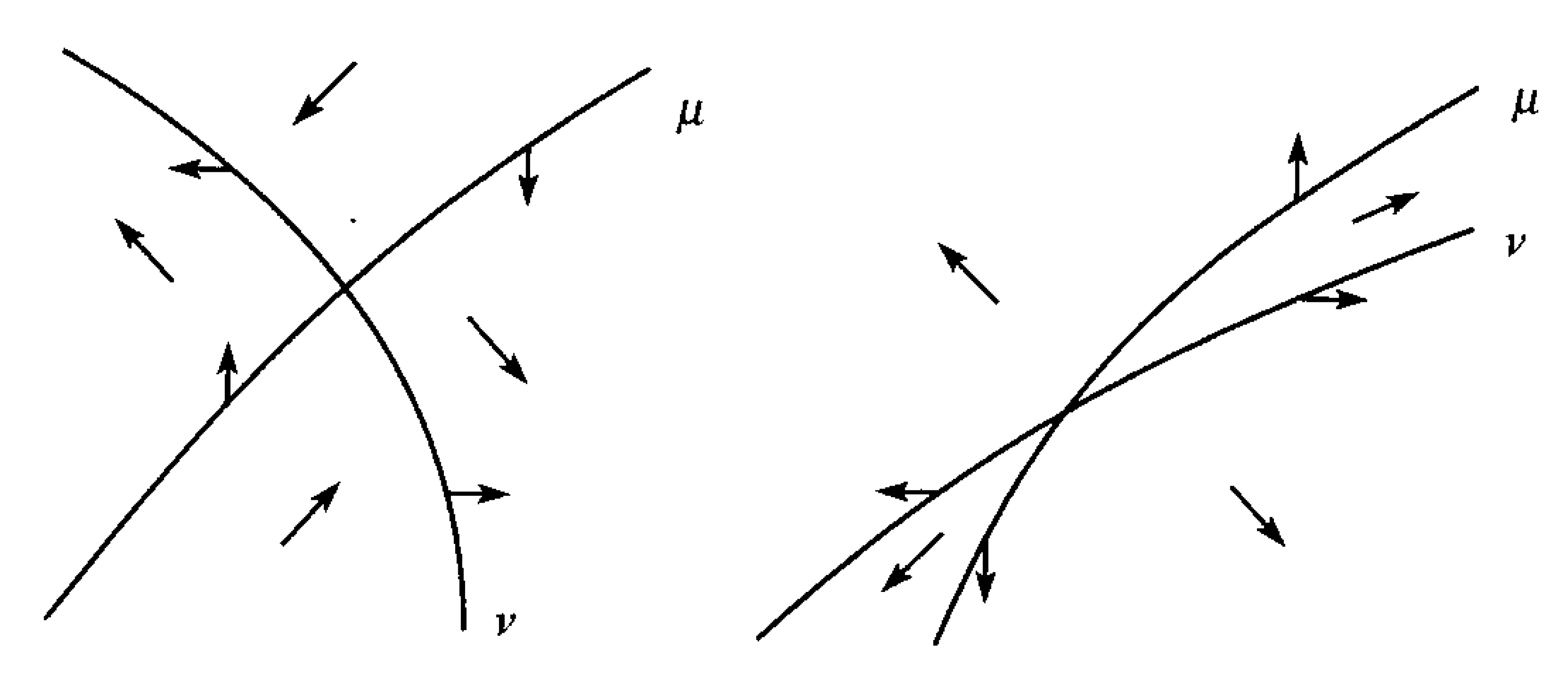


图 11.11　在顶点附近的向量场结构

由这个命题可以推出，每个基本区域和它的闭包要么是正向不变的，要么是负向不变的. 这个系统的 ω 极限点可能是什么呢？我们断言：没有闭轨. 如果有的话，一条闭轨必定包含在一个基本区域中；但沿任何解曲线，$x(t), y(t)$ 在一个基本区域中都是单调的，因而不可能有闭轨. 从而所有的 ω 极限点都是平衡点.

我们还注意到，每个解对所有的 $t \geqslant 0$ 都有定义. 这是因为每个点都位于一个大矩形 Γ 中，Γ 的四个顶点为 $(0,0),(x_0,0),(0,y_0),(x_0,y_0)$，其中 $x_0 > a, y_0 > b$；而这样的一个矩形是正向不变的 (见图 11.12). 于是我们就证明了：

定理　竞争物种系统的流 ϕ_t 具有如下的性质：对所有的点 (x,y)，其中 $x \geqslant 0, y \geqslant 0$，极限

$$\lim_{t\to\infty} \phi_t(x,y)$$

存在，且为系统的有限个平衡点之一. ■

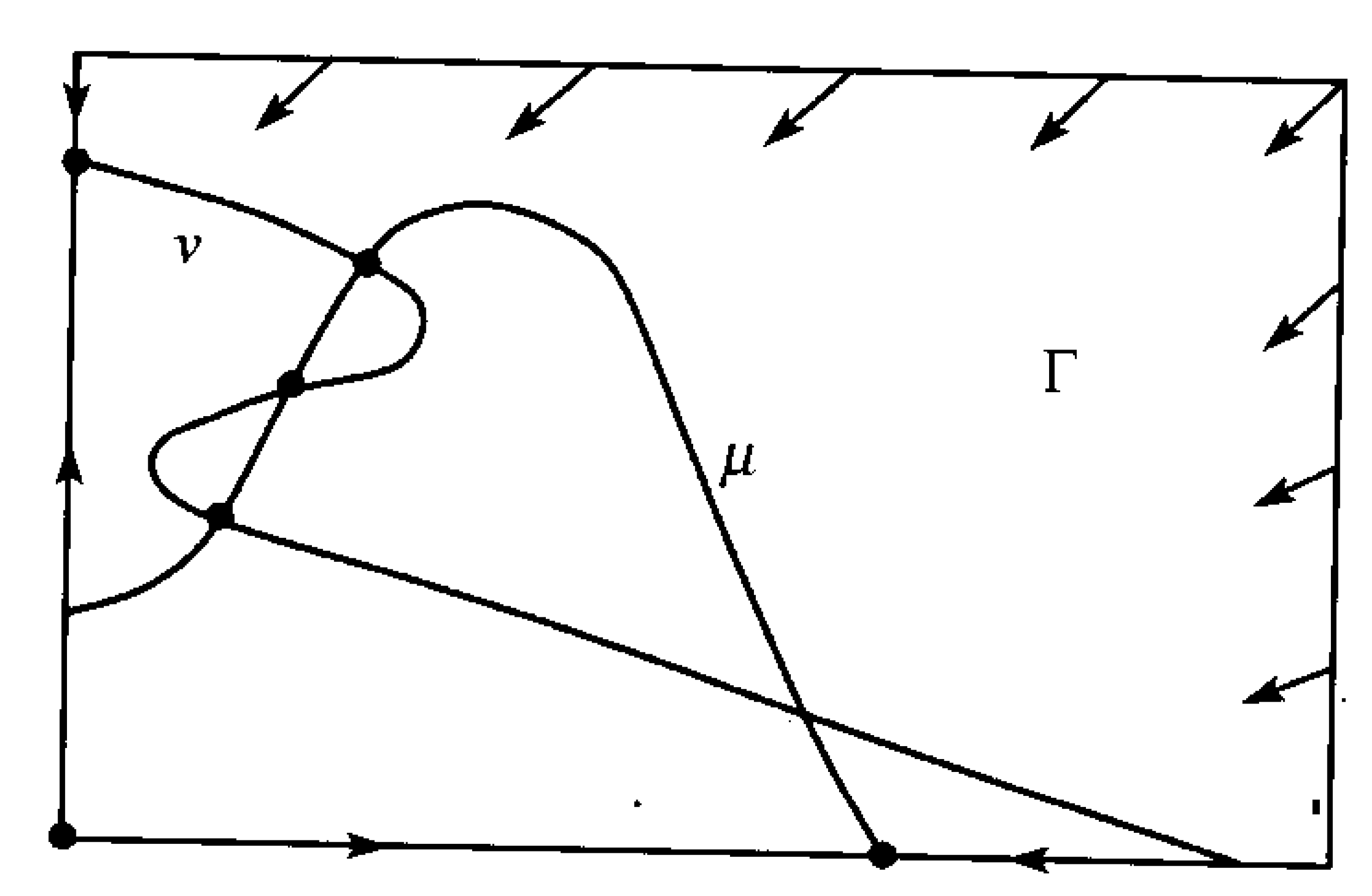


图 11.12　所有的解一定进入 Γ，然后始终逗留在 Γ 中

这样，我们就得出，两个竞争物种的数量总是趋于有限的极限物种量之一.

为了考察这些平衡点的稳定性，我们可以找出如下的结果. 如果 μ, ν 在某个顶点处的斜率都是负的，且 μ 更陡些，则这个顶点是渐近稳定的 (见图 11.13). 为

了看出这一点, 我们作一个包围该平衡点的小矩形, 使得矩形的四边都平行于坐标轴, 而且在 4 个邻接的基本区域中各有矩形的一个顶点. 这样一个矩形是正向不变的; 由于这个矩形可以取任意小, 因而这个平衡点是渐近稳定的.

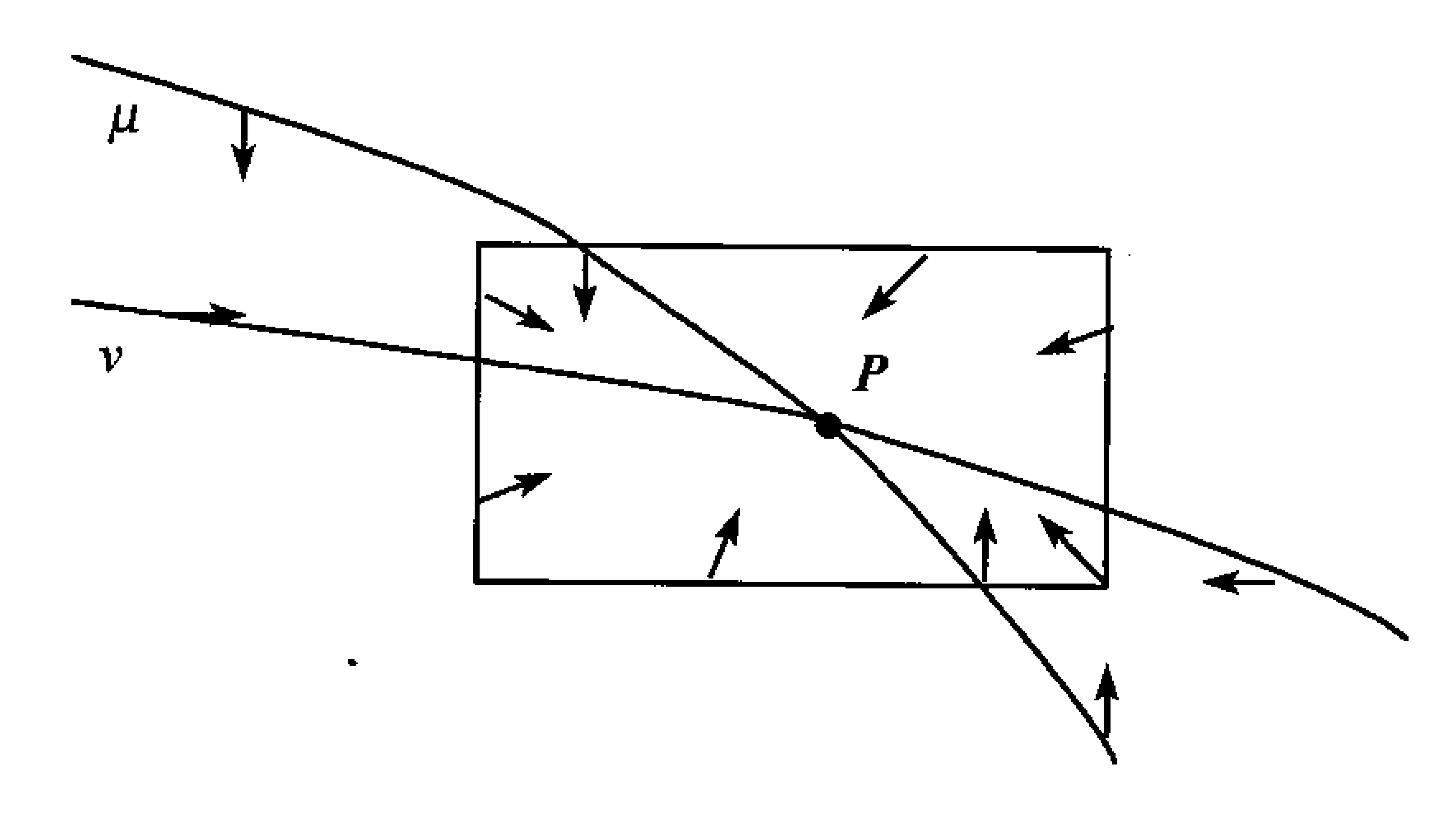


图 11.13 μ, ν 的这种结构导致一个渐近稳定的平衡点

这点还可以通过如下的方式看出. 我们有

$$\mu \text{ 的斜率 } = -\frac{M_x}{M_y} < \nu \text{ 的斜率 } = -\frac{N_x}{N_y} < 0,$$

其中 $M_x = \partial M/\partial x, M_y = \partial M/\partial y$ 在这个平衡点处的取值, 等等. 现在, 由前面的假设条件有 $M_y < 0, N_x < 0$, 从而在这个平衡点处, 我们就有 $M_x < 0, N_y < 0$. 在这个平衡点处线性化就得到矩阵

$$\begin{pmatrix} xM_x & xM_y \\ yN_x & yN_y \end{pmatrix}.$$

这个矩阵的迹为 $xM_x + yN_y < 0$, 而矩阵的行列式为 $xy(M_xN_y - M_yN_x) > 0$. 从而两个特征值都有负的实部, 从而这个平衡点是一个汇点.

通过研究 μ 和 ν 的各种不同的相交方式, 可以得出, 这个模型中仅有的其他的渐近稳定的平衡点是 $(0, b)$, 而且此时 $(0, b)$ 必须在 μ 的上方; 或是 $(a, 0)$, 但此时 $(a, 0)$ 必须在 ν 的右边. 所有其他的平衡点都是不稳定的. 一定至少有一个渐近稳定的平衡点. 如果 $(0, b)$ 不是的, 则它在 μ 的下方; 如果 $(a, 0)$ 也不是的, 则它在 ν 的左边. 这种情形, μ, ν 一定相交, 而且在 $(a, 0)$ 左边的第一个交点就是渐近稳定的.

例如, 这种分析就告诉我们, 在图 11.14 中, 只有 $\boldsymbol{P}$ 和 $(0, b)$ 是渐近稳定的; 所有其他的平衡点都是不稳定的. 特别地, 如果假设图 11.14 中的平衡点 $\boldsymbol{Q}$ 是双曲的, 则它一定是一个鞍点, 因为附近的某些解趋于它而其他的解远离它. $\boldsymbol{Z}$ 点位于 $\boldsymbol{Q}$ 的稳定曲线的一支上. 所有区域 B_∞ 中位于 $\boldsymbol{Z}$ 左边的点都趋于平衡点 $(0, b)$, 而这个区域在 $\boldsymbol{Z}$ 右边的点都趋于 $\boldsymbol{P}$. 因而当我们越过稳定曲线含 $\boldsymbol{Z}$ 点的这支时, 解

的极限行为会发生根本的变化. 由于靠 Z 右边一点点的解趋向平衡点 P, 从而在这种情形, 物种的数量趋于稳定. 另一方面, 在 Z 左边一点点的解趋于一个平衡点, 其 x 坐标为 0. 从而, 在这种情形, 其中一个物种将逐渐灭绝. 初始条件的一点小变化就导致物种命运发生戏剧性的变化. 生态学上, 这种小变化可能是因为引入一种新的杀虫剂, 迁入其中一个物种的另外成员, 发生一场森林火灾, 等等. 数学上, 这个事件就是从 P 的吸引盆跳到 $(0, b)$ 的吸引盆.

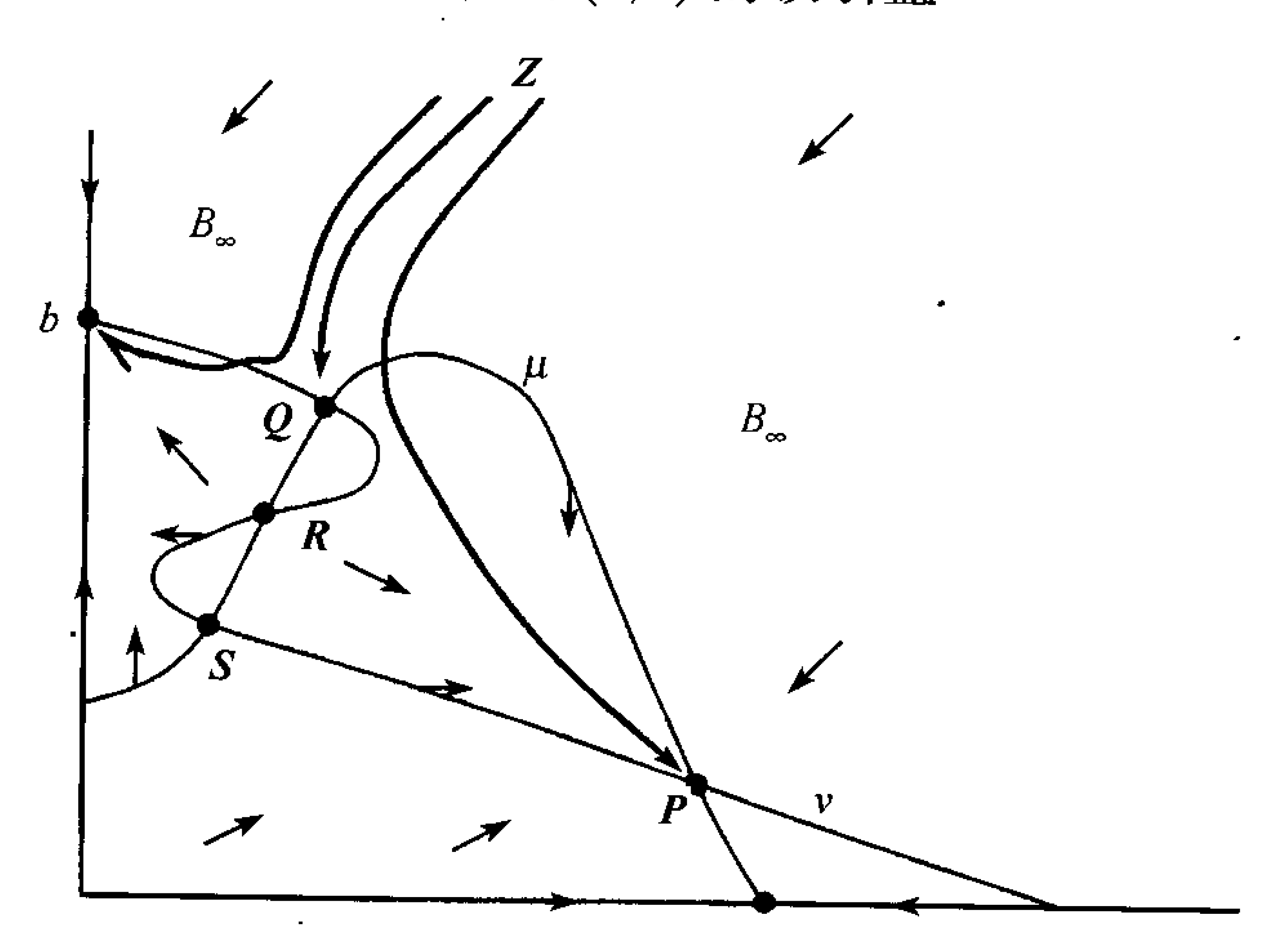


图 11.14　注意, 位于 Q 的稳定曲线上点 Z 两边的解具有非常不一样的命运

11.4　探索: 竞争与收割

在这个探索中, 我们将研究竞争物种模型, 但这里我们容许其中一个物种被收割 (迁出) 或迁入. 考虑系统

$$x' = x(1 - ax - y),$$
$$y' = y(b - x - y) + h.$$

这里 a, b, h 都是参数. 假设 $a, b > 0$. 如果 $h < 0$, 这表示在以常值速率收割物种 y, 而如果 $h > 0$, 这表示以常值速率补充物种 y. 这个探索的目的是对这些参数的所有可能取值, 完全了解这个系统. 我们照常只考虑区域 $x, y \geqslant 0$. 如果对某一 $t > 0$, 有 $y(t) < 0$, 则我们认为这个物种已经灭绝.

(1) 首先假设 $h = 0$. 通过在 ab 参数平面上作出你找到的系统的不同行为, 给出该系统的行为的一个完整大纲.

(2) 当 $h = 0$ 时, 在 ab 平面上确定发生分岔的那些点和曲线, 并描述这些分岔.

(3) 现在令 $h < 0$. 对于不同的 (固定的)h 值, 描述 ab 参数平面.

(4) 对 $h > 0$ 重复前面的实验.

(5) 描述全部三维参数空间, 你可以用任何找到的最合适的方式, 如图形、活动书页、3D 模型、动画, 等等.

习　题

1. 在 SIRS 模型中, 证明：如果物种总量不超过这种疾病的阈值, 则三角形区域 Δ 中的所有解都趋于平衡点 $(\tau, 0)$.
2. 简略地作出下面变形的捕食者/猎物系统的相图：

$$x' = x(1-x) - xy,$$

$$y' = y\left(1 - \frac{y}{x}\right).$$

3. 一种修正的捕食者/猎物系统如下：

$$x' = x(1-x) - \frac{axy}{x+1},$$

$$y' = y(1-y),$$

其中 $a > 0$ 为参数.

(a) 找出所有的平衡点并将它们分类.

(b) 对于不同的 a 值, 简略地作出零点集和相图.

(c) 描述当 a 变化时出现的所有分岔.

4. 另一种修正的捕食者/猎物系统如下：

$$x' = x(1-x) - \frac{xy}{x+b},$$

$$y' = y(1-y),$$

其中 $b > 0$ 为参数.

(a) 找出所有的平衡点并将它们分类.

(b) 对于不同的 b 值, 简略地作出零点集和相图.

(c) 描述当 b 变化时出现的所有分岔.

5. 方程组

$$x' = x(2 - x - y),$$

$$y' = y(3 - 2x - y)$$

满足 11.3 节中关于竞争种群的条件 (1) 到条件 (3). 确定该系统的相图. 解释为什么这个方程组从数学上使得两个物种同时存活成为可能, 但实际上却非常靠不住.

6. 考虑竞争种群模型

$$x' = x(a - x - ay),$$
$$y' = y(b - bx - y),$$

其中参数 a, b 是正的.

(a) 找出该系统的所有平衡点并确定其稳定类型. 这些类型当然将依赖于 a, b.

(b) 对于 a, b 的不同取值, 利用零点集确定相应的各种相图.

(c) 确定该系统发生分岔时的 a, b 值, 并描述这些分岔.

(d) 在 ab 平面上作图记录你的发现, 而且要在这个平面的每个开区域中标出相应相图的定性结构.

7. 两个物种 x, y 是**共生的**, 如果其中一个物种的数量增加会导致另一个物种的增长率增加. 因而我们假设

$$x' = M(x, y)x,$$
$$y' = N(x, y)y,$$

其中

$$\frac{\partial M}{\partial y} > 0 \quad 且 \quad \frac{\partial N}{\partial x} > 0,$$

而且 $x, y \geqslant 0$. 我们假设供给的食物总量是有限的; 从而存在 $A > 0, B > 0$ 使得

$$如果\ x > A, \quad M(x, y) < 0,$$
$$如果\ y > B, \quad N(x, y) < 0.$$

如果两个物种的数量都很小, 则它们都增加; 于是

$$M(0, 0) > 0 \quad 且 \quad N(0, 0) > 0.$$

假设曲线 $M^{-1}(0), N^{-1}$ 只有有限个交点, 而且每个交点都是横截交点. 证明下面的几条:

(a) 在区域 $0 < x < A, 0 < y < B$ 中, 所有的解都趋于一个平衡点.

(b) 没有源点.

(c) 至少有一个汇点.

(d) 如果 $\partial M/\partial x < 0$ 且 $\partial N/\partial y < 0$, 则有唯一的汇点 Z, 而且 Z 是所有满足 $x > 0, y > 0$ 的点 (x, y) 的 ω 极限集.

8. 给出一对相互毁灭物种的一个微分方程系统. 然后证明, 在合理的假设下, 两个相互毁灭的物种最后不能共存.

9. 记 y, x 分别为捕食者和猎物的数量. 设

$$x' = M(x,y)x,$$
$$y' = N(x,y)y,$$

其中 M, N 满足下面的条件.

(a) 如果没有充足的猎物, 捕食者将减少. 从而, 存在 $b > 0$, 使得当 $x < b$ 时, $N(x,y) < 0$.

(b) 猎物的增加提高捕食者的增长率, 于是 $\partial N/\partial x > 0$.

(c) 在没有捕食者的情况下, 小的猎物数量将增加, 于是 $M(0,0) > 0$.

(d) 在超过某个数量之后, 猎物的数量必定减少, 于是存在 $A > 0$, 使得当 $x > A$ 时, $M(x,y) < 0$.

(e) 捕食者的增加将导致猎物增长率的下降, 于是 $\partial M/\partial y < 0$.

(f) 两条曲线 $M^{-1}(0), N^{-1}(0)$ 横截相交, 且只交于有限个点.

证明：如果存在某一 (u,v) 使得 $M(u,v) > 0$ 且 $N(u,v) > 0$, 则要么有一个渐近稳定的平衡点, 要么有一个极限环. 进一步证明：如果极限环的个数有限而且是正的, 则其中必有一个极限环使得它的两边都有轨线盘旋地趋向它.

10. 考虑下面修正的捕食者/猎物方程：

$$x' = x(1-x) - \frac{axy}{x+c},$$
$$y' = by\left(1 - \frac{y}{x}\right),$$

其中 a, b, c 都是正的常数. 确定参数空间中的区域, 使得参数在该区域中的系统有一个稳定的平衡点, 而且该平衡点的 x, y 坐标都非零. 证明：如果这个平衡点是不稳定的, 则该系统有一个稳定的极限环.

第 12 章　电路理论中的应用

在本章, 我们首先引入一个简单但又非常基本的电子电路, 然后导出控制这个电路的微分方程. 在后面的两节, 我们将利用在第 8~10 章发展的技巧来分析这些方程的某些特殊情形; 这些情形就是经典的里纳德方程和范德波方程. 特别地, 范德波方程也许可以认为是非线性常微分方程的基本例子之一. 它有一个振动或周期解, 而且这是一个周期吸引子. 每个非平凡解都趋向这个周期解; 没有线性系统具有这一性质. 有时, 系统中的渐近稳定平衡点意味着死, 而吸引的振动解则意味着生. 在 12.4 节, 我们将给出一个从一种情形连续过渡到另一种情形的例子.

12.1　RLC 电路

在本节中, 我们引入电子电路的第一个例子. 这个例子就是图 12.1 中所示的简单而又基本的 RLC 电路. 我们从解释这个图的数学含义开始. 这个电路由 3 个**支路** 构成: 一个为电阻 R, 一个为电感 L, 以及一个电容 C. 我们认为该电路的每个支路都是某种有两个接线端的电子器件. 例如, 在这个电路中, 支路 R 的端点为 α, β, 而且所有这些端点都用电线接起来从而形成**节点**α, β, γ.

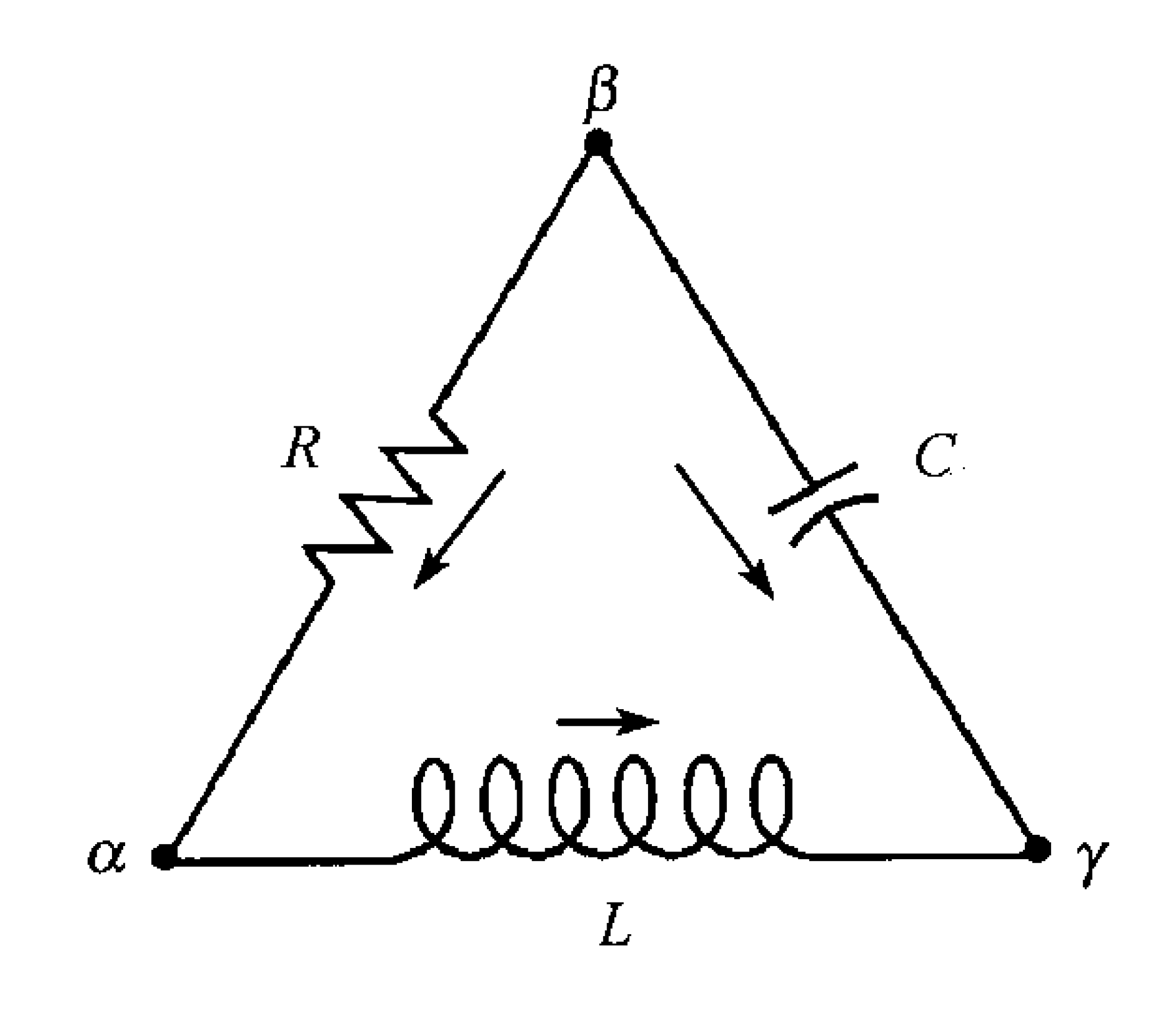


图 12.1　RLC 电路

在这个电路中, 有电流流过每个支路, 我们用一个实数来度量电流. 准确地说, 电路中的电流由 3 个实数 $i_{\mathrm{R}}, i_{\mathrm{L}}, i_{\mathrm{C}}$ 给出, 其中 i_{R} 代表经过电阻的电流, 等等. 电路中的电流类似于水管中流动的水; 对水的对应度量应该是单位时间内流过的总量, 或水经过管子中一个固定点的速度. 图中给出的支路定向的箭头告诉我们电流流动的方式 (对照水流的情形): 例如, 如果 i_{R} 是正的, 则根据箭头, 电流从 β 到 α 流过电阻 (箭头一旦在开始选定就一直不变).

于是, 在一个给定时刻, 电路中电流的状态就由 $\mathbb{R}^3$ 中的一个点 $i = (i_{\mathrm{R}}, i_{\mathrm{L}}, i_{\mathrm{C}})$ 给出. 但是**基尔霍夫电流定律** (KCL)则说, 实际中对 i 有很强的限制. KCL 声称, 流入一个节点的总电流等于流出这个节点的总电流 (对照水流情形这是显而易见

的). 对电路, 这等价于

$$\text{KCL}: \quad i_{\mathrm{R}} = i_{\mathrm{L}} = -i_{\mathrm{C}}.$$

这就定义了 $\mathbb{R}^3$ 的一个一维子空间 K_1, 也就是**物理电流的状态**空间.

我们在电容支路上选取的方向似乎不是很自然. 事实上, 这些方向的选取是任意的; 在这个例子中如此选取方向是为了使最终得到的方程直接反映这个学科的历史.

电路的**状态**由电流 $i = (i_{\mathrm{R}}, i_{\mathrm{L}}, i_{\mathrm{C}})$ 与每个支路两端的电压 (电压降落) 刻画. 分别用 $v_{\mathrm{R}}, v_{\mathrm{L}}, v_{\mathrm{C}}$ 记电阻支路、电感支路、电容支路的电压. 在水流情形, 可以认为电压降落就是水管两端的压强之差. 为了测量电压, 在每个节点 α, β, γ 放上一个伏特计 (想象成一个水压表), 读出 α 处的 $V(\alpha)$, 等等. 则 v_{R} 就是在 α, β 处的读数之差

$$V(\beta) - V(\alpha) = v_{\mathrm{R}}.$$

箭头的方向告诉我们 $v_{\mathrm{R}} = V(\beta) - V(\alpha)$ 而不是 $V(\alpha) - V(\beta)$.

电路的**自由电压状态** 就是一个点 $v = (v_{\mathrm{R}}, v_{\mathrm{L}}, v_{\mathrm{C}}) \in \mathbb{R}^3$. 此时**基尔霍夫电压定律**就对 v 给出了一个物理限制：

$$\text{KVL}: \quad v_{\mathrm{R}} + v_{\mathrm{L}} - v_{\mathrm{C}} = 0.$$

这定义了 $\mathbb{R}^3$ 中的一个二维子空间 K_2. KVL 可从 $v_{\mathrm{R}}, v_{\mathrm{L}}, v_{\mathrm{C}}$ 的伏特计读数定义中立刻得出, 即

$$v_{\mathrm{R}} + v_{\mathrm{L}} - v_{\mathrm{C}} = (V(\beta) - V(\alpha)) + (V(\alpha) - V(\gamma)) - (V(\beta) - V(\gamma)) = 0.$$

乘积空间 $\mathbb{R}^3 \times \mathbb{R}^3$ 称为电路的**状态空间**. 那些满足基尔霍夫定律的状态 (i, v) 形成了状态空间的一个三维子空间.

现在我们给出电路中的三种电子器件的数学定义. 首先考虑电阻元. R 支路上的一个电阻在 i_{R} 和 v_{R} 之间加上了一个“函数关系”. 在这个例子中, 我们认为这个关系由一个函数 $f: \mathbb{R} \to \mathbb{R}$ 定义, 且使得 $v_{\mathrm{R}} = f(i_{\mathrm{R}})$. 如果 R 是通常的线性电阻, 则 f 是线性的, 从而 $v_{\mathrm{R}} = k i_{\mathrm{R}}$. 这个关系就是著名的**欧姆定律**. 非线性函数就导致一个推广的欧姆定律. f 的图像称为这个电阻的**特性**. 一对特性的例子如图 12.2 所示. (如图 12.2b 中的特性出现在一种“隧道二极管”中.)

一个**物理状态**$(i, v) \in \mathbb{R}^3 \times \mathbb{R}^3$ 就是满足 KCL, KVL, 以及 $f(i_{\mathrm{R}}) = v_{\mathrm{R}}$ 的一个点. 这些条件就定义了**物理状态集**$\Sigma \subset \mathbb{R}^3 \times \mathbb{R}^3$. 于是, Σ 就是满足如下 3 条的 $\mathbb{R}^3 \times \mathbb{R}^3$ 中的点 $(i_{\mathrm{R}}, i_{\mathrm{L}}, i_{\mathrm{C}}, v_{\mathrm{R}}, v_{\mathrm{L}}, v_{\mathrm{C}})$ 的全体构成的集合：

(1) $i_{\mathrm{R}} = i_{\mathrm{L}} = -i_{\mathrm{C}}$ (KCL);

(2) $v_R + v_L - v_C = 0$ (KVL);

(3) $f(i_R) = v_R$ (推广的欧姆定律).

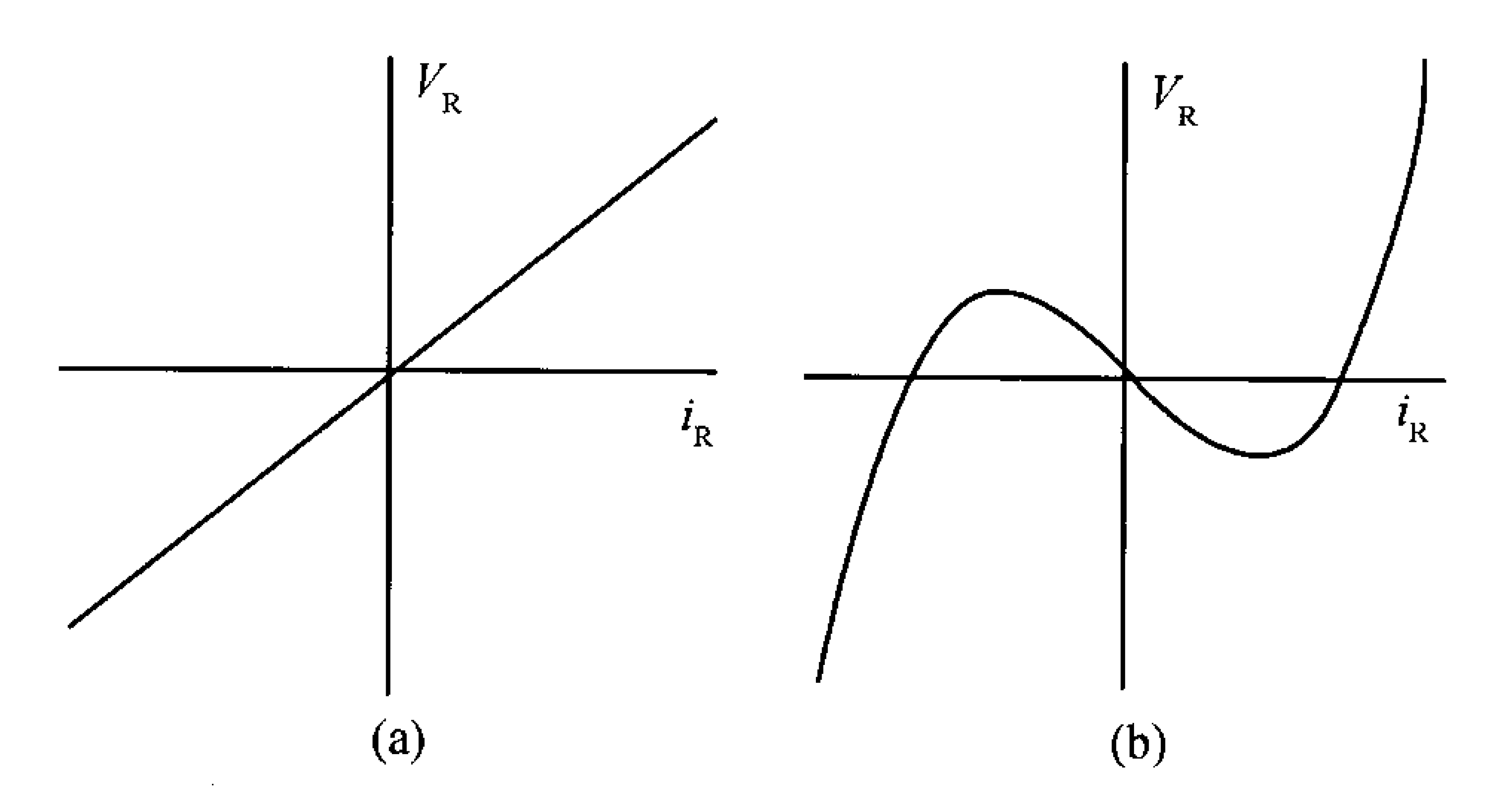


图 12.2 电阻的几种可能特性

现在我们来看控制电路的微分方程. 电感 (我们将它看成一个螺线管; 很难找到它的水流对应) 满足

$$L\frac{\mathrm{d}i_L(t)}{\mathrm{d}t} = v_L(t) \quad (\text{法拉第定律}),$$

其中 L 是一个正常数, 称为**电感**. 另一方面, 电容 (我们可以将它看成是被某种绝缘体隔开的两个金属片; 它对应于水流模型中的水槽) 满足条件:

$$C\frac{\mathrm{d}v_C(t)}{\mathrm{d}t} = i_C(t),$$

其中 C 是一个正常数, 称为**电容**.

我们将所有前面的讨论总结如下: 电路的状态由 6 个数 $(i_R, i_L, i_C, v_R, v_L, v_C)$ 给出, 这也就是状态空间 $\mathbb{R}^3 \times \mathbb{R}^3$ 中的一个点. 这些数要受到 3 个限制: 基尔霍夫电流定律、基尔霍夫电压定律以及这个电阻的特性或推广的欧姆定律. 于是, 物理状态集为某个子集 $\Sigma \subset \mathbb{R}^3 \times \mathbb{R}^3$. 一个状态随时间的变化方式就由前 2 个微分方程所决定.

下面, 观察到 i_L, v_C 可以确定其他 4 个坐标, 我们可以将物理状态集 Σ 化简. 由 KCL 有, $i_R = i_L$, $i_C = -i_L$; 由推广的欧姆定律有, $v_R = f(i_R) = f(i_L)$; 再由 KVL 有, $v_L = v_C - v_R = v_C - f(i_L)$. 因而, 我们可用 $\mathbb{R}^2$ 作为状态空间, 其中的坐标由 (i_L, v_C) 给出. 形式上, 我们定义一个映射 $\pi : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^2$, 它将 $(i, v) \in \mathbb{R}^3 \times \mathbb{R}^3$ 映到 (i_L, v_C). 然后, 我们令 $\pi_0 = \pi|\Sigma$, 即 π 在 Σ 上的限制. $\pi_0 : \Sigma \to \mathbb{R}^2$ 是一一的且满的; 它的逆由映射 $\psi : \mathbb{R}^2 \to \Sigma$ 给出, 其中

$$\psi(i_L, v_C) = (i_L, i_L, -i_L, f(i_L), v_C - f(i_L), v_C).$$

容易验证 $\psi(i_{\rm L}, v_{\rm C})$ 满足 KCL, KVL, 以及推广的欧姆定律, 因而 ψ 的确将 $\mathbb{R}^2$ 映入 Σ. 也容易看出 π_0 和 ψ 是互逆的.

于是, 我们采用 $\mathbb{R}^2$ 作为我们的状态空间. 控制状态变化的微分方程必须用现在的新坐标 $(i_{\rm L}, v_{\rm C})$ 重写:

$$
\begin{aligned}
L\frac{\mathrm{d}i_{\rm L}}{\mathrm{d}t} &= v_{\rm L} = v_{\rm C} - f(i_{\rm L}),\\
C\frac{\mathrm{d}v_{\rm C}}{\mathrm{d}t} &= i_{\rm C} = -i_{\rm L}.
\end{aligned}
$$

为了简单起见, 而且也因为这仅仅只是一个例子, 我们令 $L = C = 1$. 如果令 $x = i_{\rm L}$, $y = v_{\rm C}$, 则我们就得到下面的平面微分方程系统:

$$
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= y - f(x),\\
\frac{\mathrm{d}y}{\mathrm{d}t} &= -x.
\end{aligned}
$$

这就是著名的**里纳德方程**的一种形式. 我们在下节分析这个系统.

12.2 里纳德方程

在本节, 我们开始研究上节中来自于电路的里纳德系统的相图, 即:

$$
\begin{aligned}
\frac{\mathrm{d}x}{\mathrm{d}t} &= y - f(x),\\
\frac{\mathrm{d}y}{\mathrm{d}t} &= -x.
\end{aligned}
$$

在 $f(x) = x^3 - x$ 的特殊情形, 这个系统称为**范德波方程**.

首先考虑最简单的情形, 即 f 为线性的情形. 假设 $f(x) = kx$, 其中 $k > 0$. 则此时里纳德系统具有形式 $\boldsymbol{Y}' = \boldsymbol{A}\boldsymbol{Y}$, 其中

$$
\boldsymbol{A} = \begin{pmatrix} -k & 1 \\ -1 & 0 \end{pmatrix}.
$$

A 的特征值为

$$
\lambda_{\pm} = \frac{-k \pm \sqrt{k^2 - 4}}{2}.
$$

由于 $\lambda_{\pm}$ 要么是负的, 要么具有负的实部, 因而位于原点的平衡点是一个汇点. 当 $k < 2$ 时, 它是一个螺线汇点. 对任何的 $k > 0$, 系统的所有解都趋于原点; 物理上, 这就是电阻的耗散效果.

注意, 我们有

$$y'' = -x' = -y + kx = -y - ky',$$

从而这个系统等价于二阶方程 $y'' + ky' + y = 0$, 在初等微分方程课程中经常会遇到这个方程.

下面, 我们考虑一个一般的特性 f. 里纳德系统有唯一的平衡点, 它由 $(0, f(0))$ 给出. 在该点线性化后就得到矩阵

$$\begin{pmatrix} -f'(0) & 1 \\ -1 & 0 \end{pmatrix},$$

它的特征值为

$$\lambda_{\pm} = \frac{1}{2}\left(-f'(0) \pm \sqrt{(f'(0))^2 - 4}\right).$$

由此我们得出, 当 $f'(0) > 0$ 时, 这个平衡点是一个汇点; 当 $f'(0) < 0$ 时, 该平衡点是一个源点. 特别地, 对于 $f(x) = x^3 - x$ 的范德波方程, 唯一的平衡点是一个源点.

为了进一步分析这个系统, 我们定义函数 $W: \mathbb{R}^2 \to \mathbb{R}$, $W(x, y) = \frac{1}{2}(x^2 + y^2)$. 于是我们有

$$\dot{W} = x(y - f(x)) + y(-x) = -xf(x).$$

特别地, 如果 f 满足当 $x > 0$ 时, $f(x) > 0$, 当 $x < 0$ 时, $f(x) < 0$, 且 $f(0) = 0$, 则 W 是整个 $\mathbb{R}^2$ 上的一个严格李雅普诺夫函数. 由此可得, 在这种情形, 所有的解都趋于位于原点的唯一平衡点.

在电路理论中, 如果一个电阻的特性包含在原点以及第一、三象限的内部所构成的集合中, 就称这个电阻是**中性的**. 于是, 在中性电阻的情形, 除了在 $x = 0$ 处外, $-xf(x)$ 总是负的, 从而所有的解都趋于原点. 因而“中性”这个词正确地描述了这样一个电路的动力学.

12.3 范德波方程

本节我们继续研究里纳德方程在 $f(x) = x^3 - x$ 时的特殊情形, 也就是范德波方程:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y - x^3 + x,$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = -x.$$

记 ϕ_t 为这个系统的流. 在这种情形, 我们可以给出一个相当完全的相图分析.

定理 范德波方程有一个非平凡的周期解, 而且其他所有的解 (位于原点的平衡点除外) 都趋向这个周期解. "这个系统振荡." ■

从上节我们知道, 原点是这个系统的唯一平衡点, 而且由于 $f'(0) < 0$, 这个平衡点是一个源点. 下一步就是来证明所有的非平衡解在某种意义下绕着该平衡点顺时针"旋转". 为了看出这点, 注意, 此时的 x 零点集由 $y = x^3 - x$ 给出, 而 y 零点集为 y 轴. 我们将这些零点集每个都分成两片:

$$\nu^+ = \{(x, y)|y > 0, x = 0\},$$
$$\nu^- = \{(x, y)|y < 0, x = 0\},$$
$$g^+ = \{(x, y)|x > 0, y = x^3 - x\},$$
$$g^- = \{(x, y)|x < 0, y = x^3 - x\}.$$

这些曲线互不相交; 加上原点后, 它们就构成了 4 个基本区域 A, B, C, D 的边界, 如图 12.3 所示.

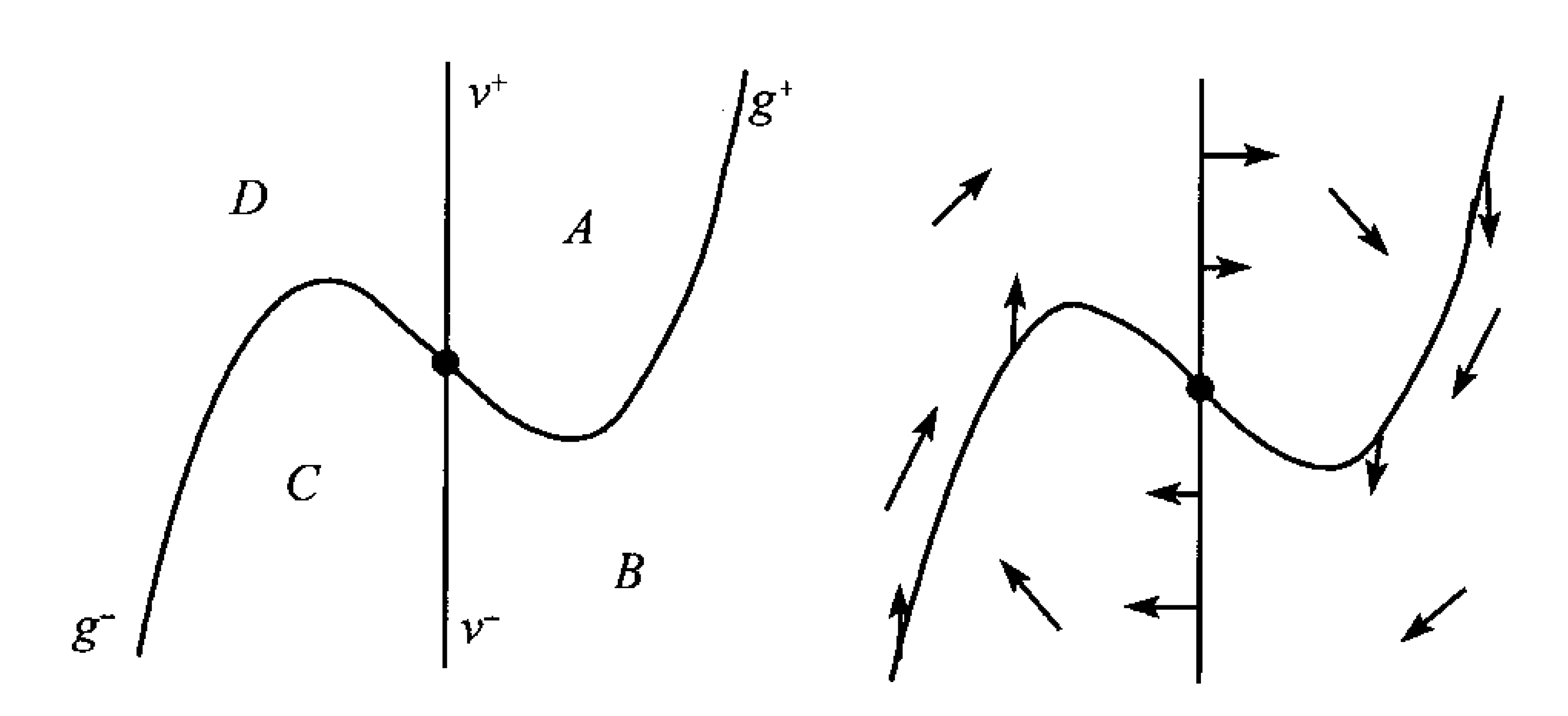


图 12.3 范德波系统的基本区域和零点集

从这个向量场在这些基本区域中的结构, 似乎可以看出所有的非平衡解都绕着原点顺时针缠绕. 的确是如此.

命题 从 ν^+ 中出发的解曲线将依次穿过 g^+, v^-, g^-, 然后再回到 v^+.

证明 因为 $x'(0) < 0$, 任何从 ν^+ 出发的解将立即进入区域 A. 在 A 中, 我们有 $y' < 0$, 从而这个解沿 y 方向必定减少. 由于这个解不能趋于源点, 从而它最终必定要与 g^+ 相交. 在 g^+ 上, 我们有 $x' = 0, y' < 0$. 从而, 这个解要穿过 g^+, 然后进入区域 B. 一旦进入 B 内部, 这个解就向西南方向前进. 由于向量场在 g^+ 上竖直向下, 这个解不能再进入区域 A. 现在有两种可能: 要么解穿过 ν^-, 要么这个解在 y 方向趋于 $-\infty$, 而且总不穿过 ν^-.

我们断言后面一种情形不会发生. 假设发生了这种情形. 令 (x_0, y_0) 是这个解上位于区域 B 中的一个点, 考虑 $\phi_t(x_0, y_0) = (x(t), y(t))$. 由于 $x(t)$ 从不取 0, 因而解曲线在所有时间都位于带形区域 $S : 0 < x \leqslant x_0, y \leqslant y_0$ 中, 而且对于某个 t_0, 当 $t \to t_0$ 时, $y(t) \to -\infty$. 我们首先可以看出, 事实上有 $t_0 = \infty$. 为了看出这一点, 注

意

$$y(t)-y_0=\int_0^t y'(s)\mathrm{d}s=\int_0^t -x(s)\mathrm{d}s.$$

但由于 $0<x(s)\leqslant x_0$, 从而只有在 $t\to\infty$ 时, 我们才可能有 $y(t)\to-\infty$.

现在考虑 $0\leqslant t<\infty$ 时的 $x(t)$. 我们有 $x'=y-x^3+x$. 由于 $-x^3+x$ 的值在带形 S 中是有界的, 而且当 $t\to\infty$ 时, $y(t)\to-\infty$, 从而当 $t\to\infty$ 时, 就有

$$x(t)-x_0=\int_0^t x'(s)\mathrm{d}s\to-\infty.$$

但这与假设 $x(t)>0$ 矛盾.

从而这个解必定穿过 ν^-. 现在, 这个向量场关于原点是反对称的, 即, 如果记 $G(x,y)$ 为范德波向量场, 则有 $G(-x,-y)=-G(x,y)$. 利用这种对称性, 解就必定以类似的方式穿过区域 C,D. ■

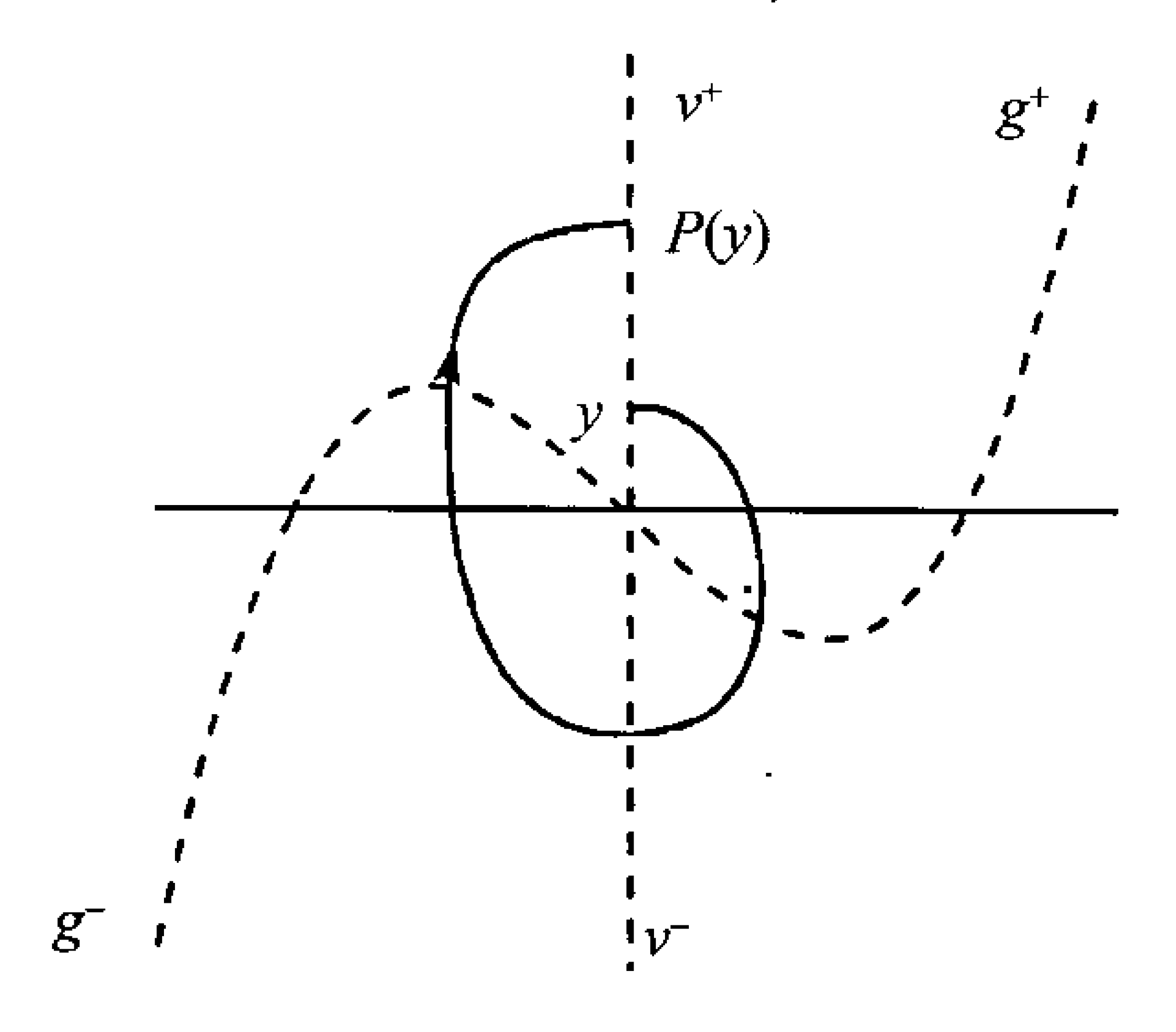


图 12.4　ν^+ 上的庞加莱映射

根据这个结果, 我们可以在半直线 ν^+ 上定义一个庞加莱映射 P. 给定 $(0,y_0)\in\nu^+$, 我们定义 $P(y_0)$ 是 $\phi_t(0,y_0)$ 在 $t>0$ 时首次返回到 ν^+ 的 y 坐标 (见图 12.4). 与 10.3 节中一样, P 是一个一一的 C^∞ 函数. 而且这个庞加莱映射还是满的. 这是因为, 与上个命题一样, 在时间向后时, 从 ν^+ 出发的解最终也要回到 ν^+. 记 $P^n=P\circ P^{n-1}$ 为 P 与自己的 n 次复合.

现在我们的目的是证明下面的定理:

定理　ν^+ 上的庞加莱映射有唯一的不动点. 而且, 对于任意非零的 $y_0\in\nu^+$, 当 $n\to\infty$ 时, 序列 $P^n(y_0)$ 都趋于这个不动点. ■

显然, P 的任一不动点都位于一个周期解上. 另一方面, 如果 $P(y_0)\neq y_0$, 则过 $(0,y_0)$ 的解不可能是周期的. 事实上, 如果 $P(y_0)>y_0$, 则像 10.4 节中一样, 可以说明, $\phi_t(0,y_0)$ 与 ν^+ 依次的交点是一个单调序列. 于是, 这个解依次穿过 ν^+ 将得到一个递增的点列, 从而这个解永远也不可能与它自己相交. $P(y_0)<y_0$ 的情形是类似的.

我们还可以定义一个"半庞加莱映射" $\alpha:\nu^+\to\nu^-$, 此时 $\alpha(y)$ 是 $\phi_t(0,y)$ 在 $t>0$ 时与 ν^- 的第一个交点的 y 坐标. 再定义

$$\delta(y)=\frac{1}{2}(\alpha(y)^2-y^2).$$

以后还需要用到如下的事实: 存在唯一的点 $(0,y^*)\in\nu^+$ 和时刻 $t^*>0$ 满足

(1) 当 $0 < t < t^*$ 时, $\phi_t(0, y^*) \in A$;

(2) $\phi_{t^*}(0, y^*) = (1, 0) \in g^+$.

见图 12.5.

这个定理可由下面相当精细的结果直接得出.

命题 函数 $\delta(y)$ 满足:

(1) 如果 $0 < y < y^*$, 则 $\delta(y) > 0$;

(2) 对于 $y > y^*$, 当 $y \to \infty$ 时, $\delta(y)$ 单调下降趋于 $-\infty$. ■

我们随后就证明这个命题; 但我们先利用它来完成定理的证明. 我们要利用这个向量场关于原点的反对称性. 这蕴涵, 如果 $(x(t), y(t))$ 是一条解曲线, 则 $(-x(t), -y(t))$ 也是一条解曲线.

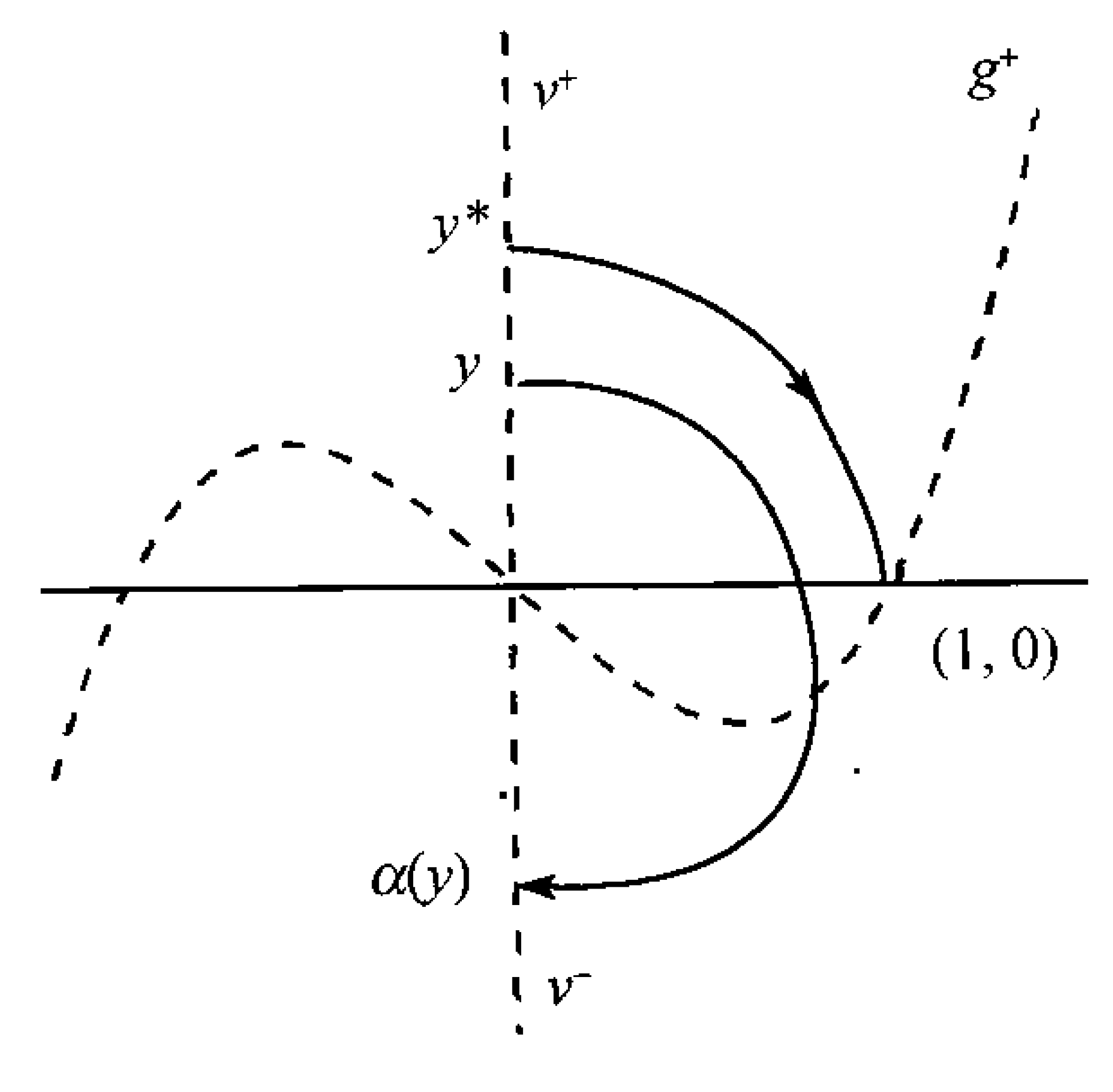


图 12.5 半庞加莱映射

$\delta(y)$ 的部分图像如图 12.6 所示. 由介值定理和上一命题可知, 存在唯一的 $y_0 \in \nu^+$ 满足 $\delta(y_0) = 0$. 于是 $\alpha(y_0) = -y_0$, 由向量场关于原点的反对称性可知, 过 $(0, y_0)$ 的解是周期的. 由于除了 y_0 外, $\delta(y) \neq 0$, 我们有, 对于其他所有的 y 值, $\alpha(y) \neq y$, 由此可得 $\phi_t(0, y_0)$ 就是唯一的周期解.

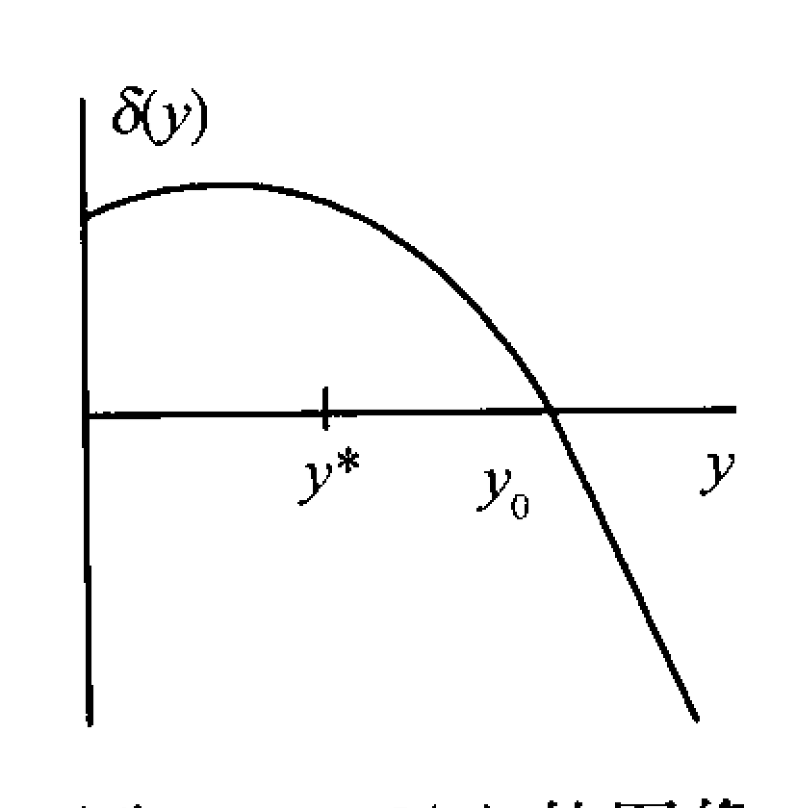


图 12.6 $\delta(y)$ 的图像

下面我们证明所有其他的解 (平衡点除外) 都趋向这个周期解. 一直到证明结束, 我们都令 $\beta : \nu^- \to \nu^+$, 它将 ν^- 上的每个点都映到相应的解 (在 $t > 0$ 时) 与 ν^+ 的第一个交点. 由对称性, 我们有

$$\beta(y) = -\alpha(-y).$$

还应注意, $P(y) = \beta \circ \alpha(y)$.

我们将 y 轴与 y 坐标中的实数等同. 于是, 如果 $y_1, y_2 \in \nu^+ \cup \nu^-$, 当 y_1 在 y_2 上方时, 我们就写成 $y_1 > y_2$. 注意, α, β 反转这个定向, 而 P 则保持这个定向.

现在取 $y \in \nu^+$ 使得 $y > y_0$. 由于 $\alpha(y_0) = -y_0$, 我们有 $\alpha(y) < -y_0$ 且 $P(y) > y_0$. 另一方面, $\delta(y) < 0$, 这蕴含 $\alpha(y) > -y$. 从而, $P(y) = \beta(\alpha(y)) < y$. 我们就证明了 $y > y_0$ 蕴涵 $y > P(y) > y_0$. 类似地, $P(y) > P(P(y)) > y_0$, 从而由归纳法可得, 对所有的 $n > 0$, 都有 $P^n(y) > P^{n+1}(y) > y_0$.

下降序列 $P^n(y)$ 在 ν^+ 中有一个极限 $y_1 \geqslant y_0$. 注意 y_1 是 P 的一个不动点, 这是因为, 根据 P 的连续性, 我们有

$$P(y_1) - y_1 = \lim_{n\to\infty} P(P^n(y)) - y_1 = y_1 - y_1 = 0.$$

由于 P 只有一个不动点, 我们有 $y_1 = y_0$. 这就说明在 $t \to \infty$ 时, 过 y 的解盘旋地趋向周期解 (见图 12.7). 同样的结论在 $y < y_0$ 时也成立; 我们将细节留给读者. 由于除平衡点外的所有解都与 ν^+ 相交, 这就完成了定理的证明. ■

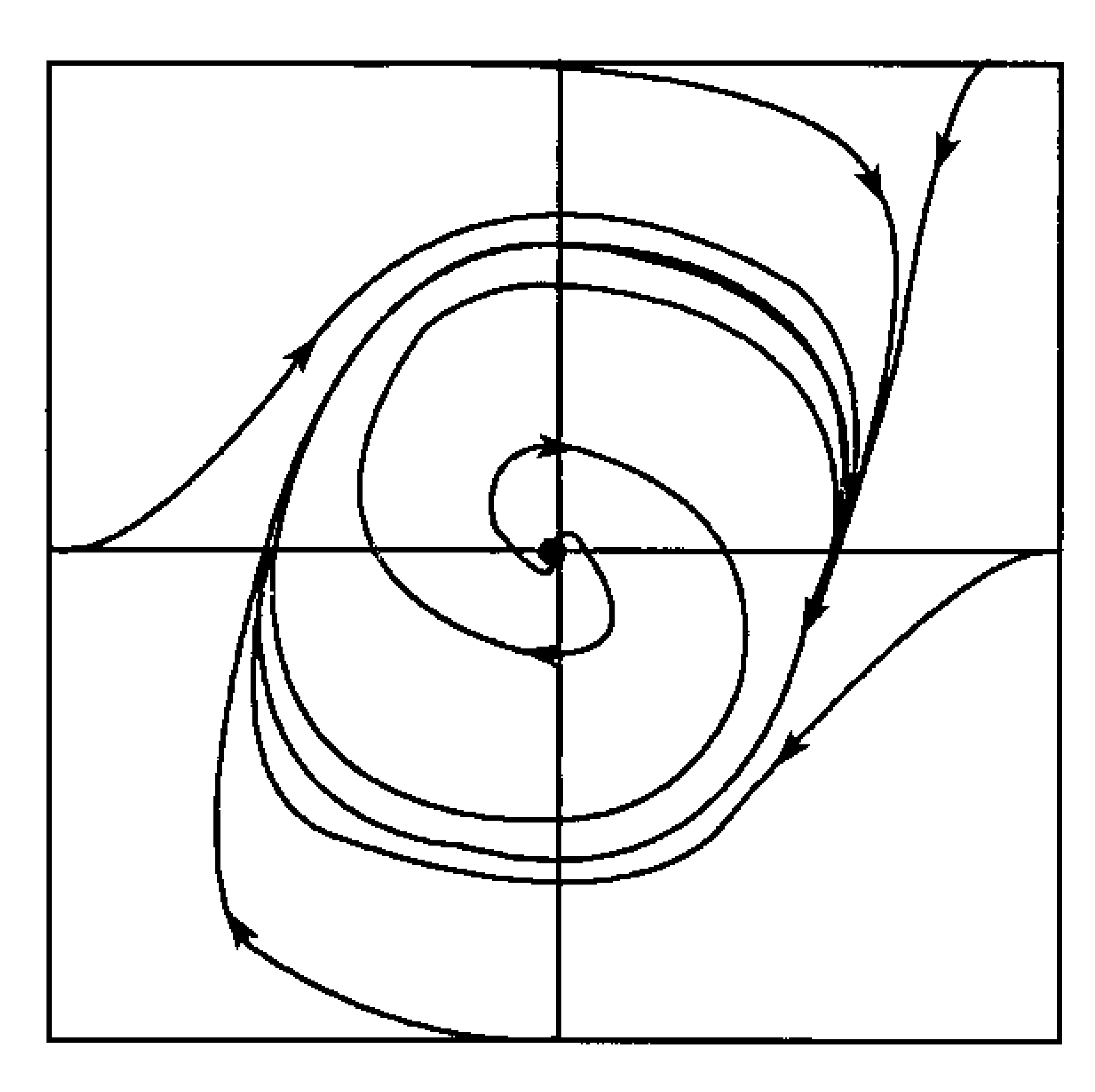


图 12.7　范德波方程的相图

最后, 我们回来证明命题. 我们采用如下的记号. 令 $\gamma : [a, b] \to \mathbb{R}^2$ 是平面上一条光滑曲线, 再令 $F : \mathbb{R}^2 \to \mathbb{R}$. 我们将 $\gamma(t)$ 写成 $\gamma(t) = (x(t), y(t))$, 而且定义

$$\int_\gamma F(x, y) = \int_a^b F(x(t), y(t))\mathrm{d}t.$$

如果对于所有的 $a \leqslant t \leqslant b$, 都有 $x'(t) \neq 0$, 则沿 γ, y 是 x 的一个函数, 于是我们可以写成 $y = y(x)$. 在这种情形, 我们可以作变量替换:

$$\int_a^b F(x(t), y(t))\mathrm{d}t = \int_{x(a)}^{x(b)} F(x, y(x))\frac{\mathrm{d}t}{\mathrm{d}x}\mathrm{d}x.$$

于是,

$$\int_\gamma F(x, y) = \int_{x(a)}^{x(b)} \frac{F(x, y(x))}{\mathrm{d}x/\mathrm{d}t}\mathrm{d}x.$$

在 $y'(t) \neq 0$ 时, 我们有类似的表达式.

现在考虑上节中引入的函数

$$W(x, y) = \frac{1}{2}(x^2 + y^2).$$

令 $p \in \nu^+$, 设 $\alpha(p) = \phi_\tau(p)$. 令 $\gamma(t) = (x(t), y(t))$, $0 \leqslant t \leqslant \tau = \tau(p)$ 为将 $p \in \nu^+$ 连到 $\alpha(p) \in \nu^-$ 的解曲线. 根据定义,

$$\delta(p) = \frac{1}{2}\left(y(\tau)^2 - y(0)^2\right) = W(x(\tau), y(\tau)) - W(x(0), y(0)).$$

从而,

$$\delta(p) = \int_0^\tau \frac{\mathrm{d}}{\mathrm{d}t} W(x(t), y(t))\mathrm{d}t.$$

在 12.2 节中我们已经有

$$\dot{W} = -xf(x) = -x(x^3 - x).$$

于是我们就有

$$\delta(p)=\int_0^\tau -x(t)(x(t)^3-x(t))\mathrm{d}t=\int_0^\tau x(t)^2(1-x(t)^2)\mathrm{d}t.$$

因为当 $0<x(t)<1$ 时, 上述积分为正, 由此可立刻得出命题中第 1 条.

我们可以将最后一个等式写成

$$\delta(p)=\int_\gamma x^2(1-x^2).$$

我们将注意力限制到 $p\in\nu^+$, $p>y^*$. 我们将相应的解曲线 γ 分成如图 12.8 所示的三条曲线 $\gamma_1,\gamma_2,\gamma_3$. 曲线 γ_1 和 γ_3 由 $0\leqslant x\leqslant 1$ 所界定, 而曲线 γ_2 则由 $y_1\leqslant y\leqslant y_2$ 所界定. 于是,

$$\delta(p)=\delta_1(p)+\delta_2(p)+\delta_3(p),$$

其中

$$\delta_i(p)=\int_{\gamma_i} x^2(1-x^2),\quad i=1,2,3.$$

注意, 沿 γ_1, $y(t)$ 可以看成是 x 的一个函数. 从而, 我们有

$$\delta_1(p)=\int_0^1\frac{x^2(1-x^2)}{\mathrm{d}x/\mathrm{d}t}\mathrm{d}x=\int_0^1\frac{x^2(1-x^2)}{y-f(x)}\mathrm{d}x,$$

其中 $f(x)=x^3-x$. 当 p 沿 y 轴向上移动, 而 (x,y) 在相应的 γ_1 上取值时, $y-f(x)$ 增加. 从而, δ_1 随着 p 的增加而减少. 类似地, δ_3 随着 p 的增加而减少.

在 γ_2 上, $x(t)$ 可以看成是 y 的一个函数, 它对 $y\in[y_1,y_2]$ 有定义, 而且 $x\geqslant 1$. 从而, 由 $\mathrm{d}y/\mathrm{d}t=-x$, 可得

$$\begin{aligned}\delta_2(p)&=\int_{y_2}^{y_1}-x(y)\big(1-x(y)^2\big)\mathrm{d}y\\&=\int_{y_1}^{y_2}x(y)\big(1-x(y)^2\big)\mathrm{d}y,\end{aligned}$$

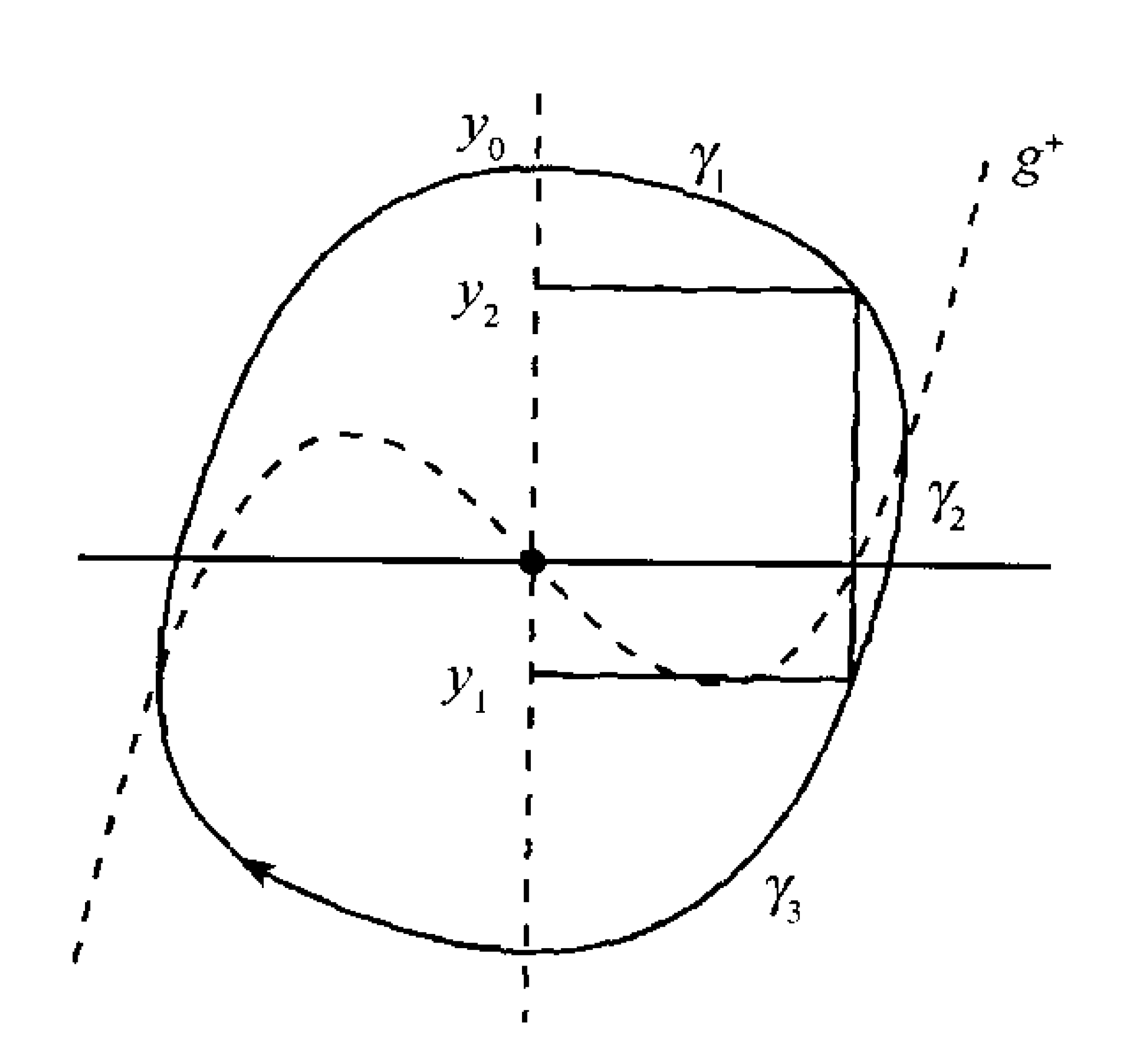


图 12.8　在过 y_0 的闭轨上所描述的曲线 $\gamma_1,\gamma_2,\gamma_3$

从而 $\delta_2(p)$ 是负的.

当 p 增加时, 积分区间 $[y_1,y_2]$ 变得越来越大. 函数 $y\to x(y)$ 依赖于 p, 因而我们将它记为 $x_p(y)$. 随着 p 的增加, 曲线 γ_2 向右边移动; 于是 $x_p(y)$ 增加, 从而 $x_p(y)(1-x_p(y)^2)$ 减少. 由此可得, $\delta_2(p)$ 随着 p 的增加而减少; 而且显然有 $\lim_{p\to\infty}\delta_2(p)=-\infty$. 于是, $\delta(p)$ 随着 p 的增加而减少, 而且在 $p\to\infty$ 时趋于 $-\infty$. 命题证毕.

12.4 一个霍普夫分岔

现在我们来描述一类更一般的电路方程, 此时电阻特性 $f_\mu(x)$ 依赖于一个参数 μ. (μ 可能是电阻的温度.) 于是, 描述电阻的物理行为的就是 $\mathbb{R}^2$ 中的微分方程系统:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y - f_\mu(x),$$
$$\frac{\mathrm{d}y}{\mathrm{d}t} = -x.$$

作为一个例子, 假设 f_μ 具有如下的特殊形式:

$$f_\mu(x) = x^3 - \mu x,$$

其中参数 μ 位于区间 $[-1,1]$ 中. 当 $\mu = 1$ 时, 我们就得到上节中的范德波系统. 像以前一样, 原点是这个系统的唯一平衡点. 该点的线性化系统为

$$\boldsymbol{Y}' = \begin{pmatrix} \mu & 1 \\ -1 & 0 \end{pmatrix} \boldsymbol{Y},$$

它的特征值是

$$\lambda_\pm = \frac{1}{2}\left(\mu \pm \sqrt{\mu^2 - 4}\right).$$

从而, 当 $-1 \leqslant \mu < 0$ 时, 原点是一个螺线汇点, 而当 $0 < \mu \leqslant 1$ 时, 原点是一个螺线源点. 事实上, 当 $-1 \leqslant \mu \leqslant 0$ 时, f_μ 的图像位于第一、三象限, 从而电阻是中性的. 因而在这种情形, 所有的解都趋于原点. 即使在 $\mu = 0$ 时, 线性化得到的是一个中心, 情况依然如此. 物理上, 这个电路是死的, 也就是说, 过了一段时间后, 所有的电流和电压都将变为 0 (或者, 和 0 要多近有多近).

然而, 当 μ 为正时, 电路就变得有生命了. 它开始振荡. 对这个系统在 μ 位于区间 $(0,1]$ 时, 运用 12.3 节的分析就可以得出这个事实. 于是, 当 μ 增加并通过 0 时, 我们就看到了一个 (唯一的) 周期解 γ_μ 的产生 (见本章的习题 4). 与前面一样, 这个解吸引所有的非零解. 同 8.5 节一样, 这是**霍普夫分岔**的一个例子. 进一步用 12.3 节中的想法进行细致的分析就可以证明, 当 μ 从 $\mu > 0$ 的方向趋于 0 时, $\gamma_\mu \to 0$. 图 12.9 显示了与这个分岔相关的几个相图.

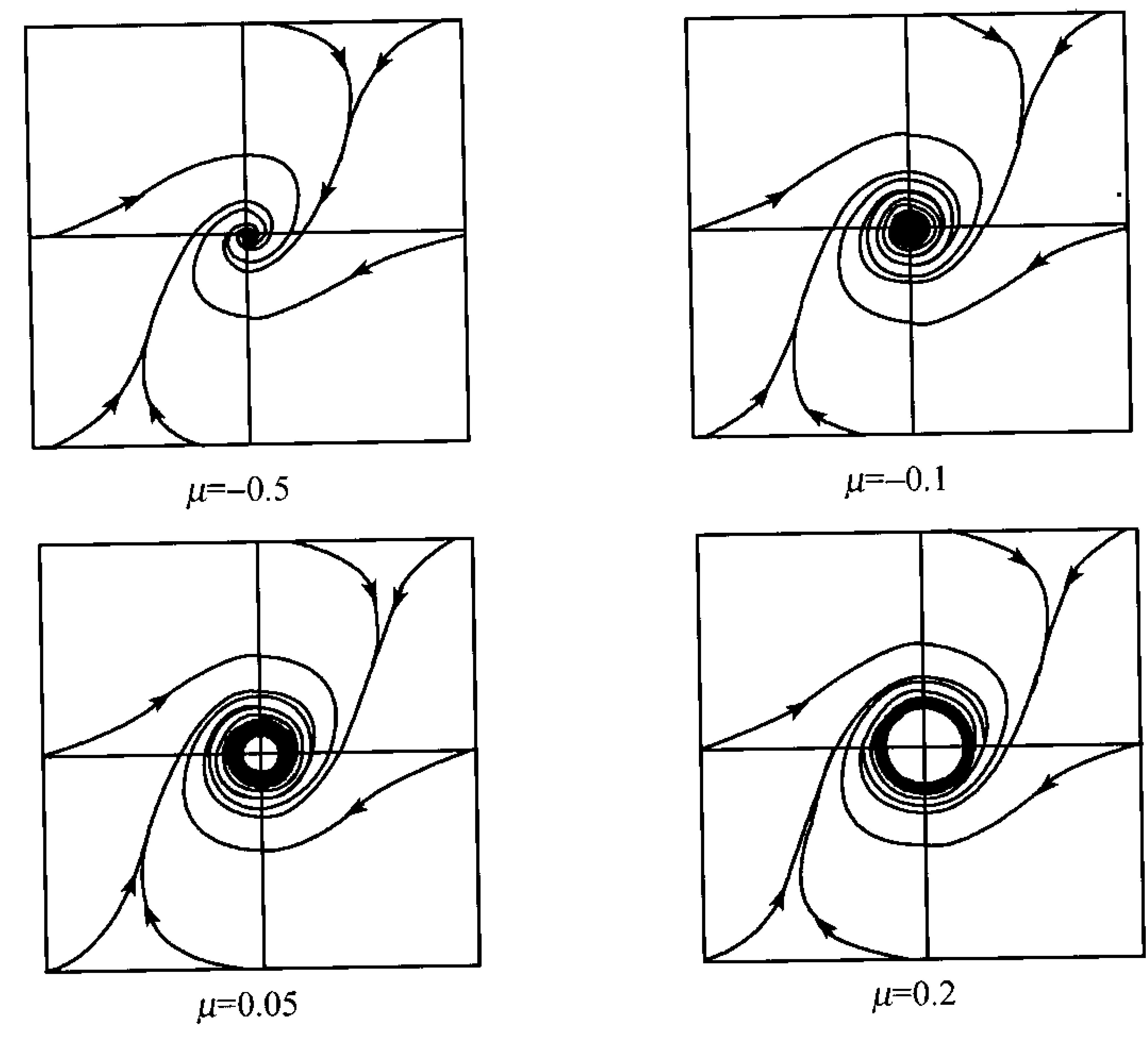


图 12.9 系统 $x' = y - x^3 + \mu x, y' = -x$ 中的霍普夫分岔

12.5 探索：神经动力学

在 20 世纪 50 年代, Hodgkin 和 Huxley 发展了一种用以研究巨型乌贼的神经冲动的模型 [23], 这是神经冲动研究中最重要的进展之一. 他们发展了一个四维的微分方程系统来描述神经信号沿细胞膜的电气化学传输, 正是因为这个工作, 他们后来获得了诺贝尔奖. 粗略地说, 这个系统类似于电子电路中产生的系统. 神经元由一个接收电刺激的细胞体构成. 这个刺激然后沿神经轴突 (它可以看成是通过一组突触连接到其他神经元的一条电缆) 传输. 当然, 这种运动不是真正的电流, 因为此时流动的不是真正的电子, 而是离子 (主要是钠离子和钾离子). 在 [15] 或 [34] 中有关于这些系统背后的神经生物学的初级介绍.

四维的 Hodgkin-Huxley 系统很难处理, 这主要是因为这些方程本质上具有很高的非线性性. 从数学上看, 一个重要的突破是由 Fitzhugh [18] 和 Nagumo 等 [35] 完成的, 他们提出了 Hodgkin-Huxley 系统的一个较简单的模型. 虽然这个系统不如原来的系统在生物学上那么准确, 但它仍然抓住了神经刺激的本质行为, 包括下面将要提到的**刺激反应**现象.

Fitzhugh-Nagumo 系统的方程为

$$x' = y + x - \frac{x^3}{3} + I, \quad y' = -x + a - by,$$

其中 a, b 为常数, 且满足

$$0 < \frac{3}{2}(1-a) < b < 1,$$

而 I 则是一个参数. 在这些方程中, x 类似于电压, 代表该系统的刺激反应; 变量 y 代表倾向于使系统静止的其他力的合力. 参数 I 是一个导致该系统兴奋的刺激参数; I 类似于一个外加电流. 注意这些方程与 12.3 节中的范德波方程的相似性.

(1) 首先假设 $I=0$. 证明这个系统有唯一的平衡点 (x_0, y_0). 提示：利用该系统的零点集的几何, 而不是显式地解出这些方程. 同时别忘了对 a, b 的限制.

(2) 证明这个平衡点总是一个汇点.

(3) 假设 $I \neq 0$. 证明仍有唯一的平衡点 (x_I, y_I), 其中 x_I 随着 I 单调地变化.

(4) 确定 x_I 的值使得这个平衡点是一个源点, 并且证明此时必定有一个稳定的极限环.

(5) 当 $I \neq 0$ 时, 点 (x_0, y_0) 不再是一个平衡点. 但是我们仍然可以考虑过该点的解. 描述当 I 从 0 离开时, 这个解的定性本质. 用数学术语解释为什么生物学家们认为这种现象是神经元的“兴奋”.

(6) 考虑 $a = I = 0$ 的特殊情形. 对每个 $b > 0$(不再限制 $b < 1$), 尽可能完备地描述相平面. 描述出现的所有分岔.

(7) 现在让 I 也变化, 再描述出现的所有分岔. 在 Ib 平面上 $(b > 0)$, 尽可能详细地描述出现的各种相图.

(8) 将上一问题中的分析扩充到 $b \leqslant 0$ 的情形.

(9) 现在固定 $b = 0$ 而让 a 和 I 变化. 在这种情形, 简略地作出分岔平面 (即 Ia 平面).

习　题

1. 找出下面微分方程的相图：

$$\begin{aligned} x' &= y - f(x), \quad f(x) = x^2, \\ y' &= -x. \end{aligned}$$

提示：利用关于 y 轴的对称性.

2. 令

$$f(x) = \begin{cases} 2x-3, & \text{如果 } x > 1, \\ -x, & \text{如果 } -1 \leqslant x \leqslant 1, \\ 2x+3, & \text{如果 } x < -1. \end{cases}$$

考虑系统

$$x' = y - f(x),$$
$$y' = -x.$$

(a) 简略地作出该系统的相图.

(b) 证明该系统有唯一的闭轨.

3. 令

$$f_a(x) = \begin{cases} 2x + a - 2, & \text{如果 } x > 1, \\ ax, & \text{如果 } -1 \leqslant x \leqslant 1, \\ 2x - a + 2, & \text{如果 } x < -1. \end{cases}$$

考虑系统

$$x' = y - f_a(x),$$
$$y' = -x.$$

(a) 对于 a 的不同取值, 简略地作出相应系统的相平面.

(b) 描述在 $a = 0$ 时出现的分岔.

4. 考虑在 12.4 节中描述的系统:

$$x' = y - (x^3 - \mu x),$$
$$y' = -x,$$

这里的参数 μ 满足 $0 \leqslant \mu < 1$. 将 "在 $\mu = 0$ 时出现一个霍普夫分岔" 这一断言的证明细节补全.

5. 考虑系统

$$x' = \mu(y - (x^3 - x)), \quad \mu > 0,$$
$$y' = -x.$$

证明这个系统具有唯一的非平凡周期解 γ_μ. 证明当 $\mu \to \infty$ 时, γ_μ 趋向一条闭曲线, 这条闭曲线由两条水平线段以及 $y = x^3 - x$ 上的两个弧段构成, 如图 12.10 所示. 一个这种类型的解称为一个松弛振动. 当 μ 很大时, 沿着相应的周期解, 有两个非常不同的时间尺度. 当沿水平方向运动时, 我们有非常大的 x', 从而这个解通过得非常快. 另一方面, 在立方零点集附近, $x' = 0$, 且 y' 是有界的. 于是, 相对来说, 这时的通过就要慢得多.

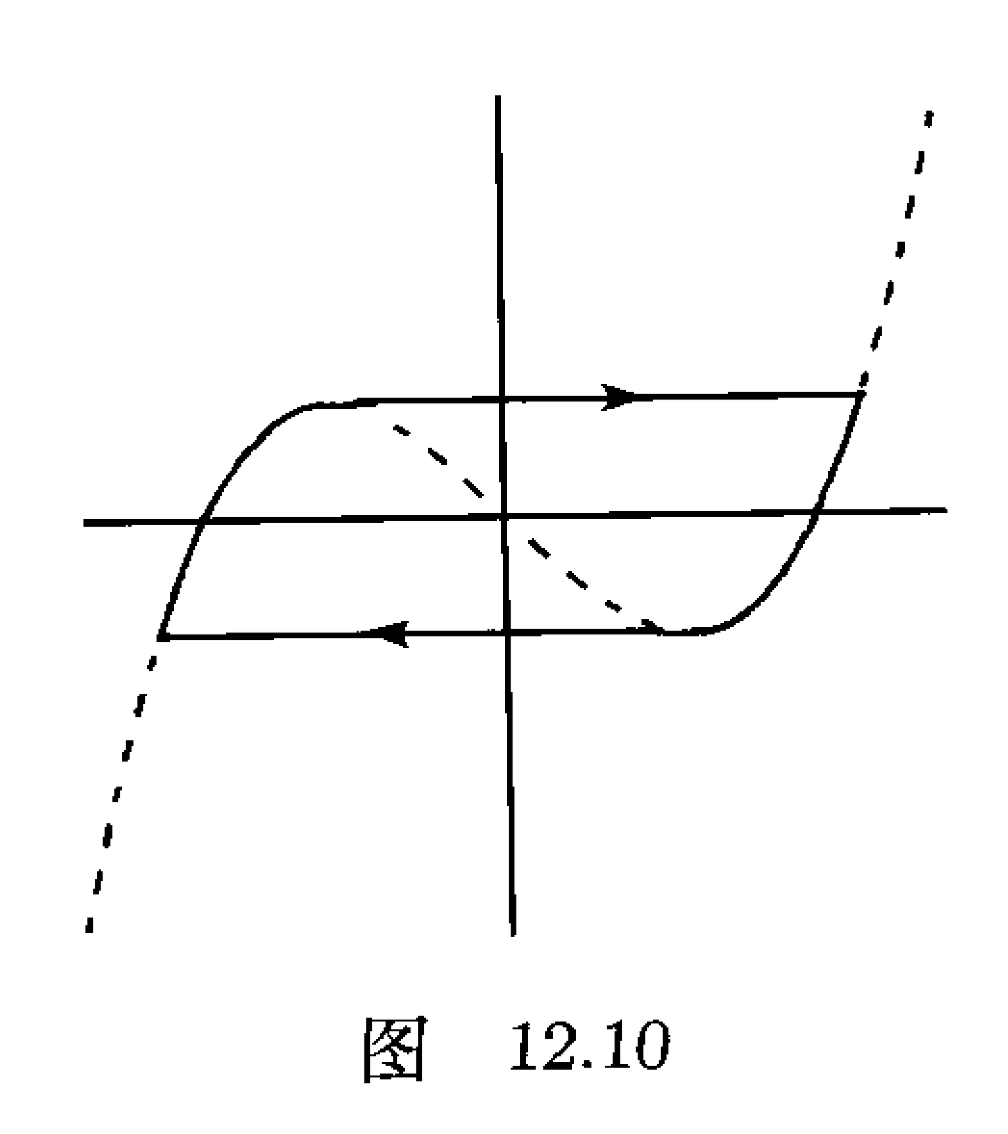


图　12.10

6. 找出图 12.11 中的电路网的微分方程, 其中电阻是电压控制的, 即, 该电阻的特性是一个函数 $g:\mathbb{R}\to\mathbb{R}, i_{\mathrm{R}}=g(v_{\mathrm{R}})$ 的图像.

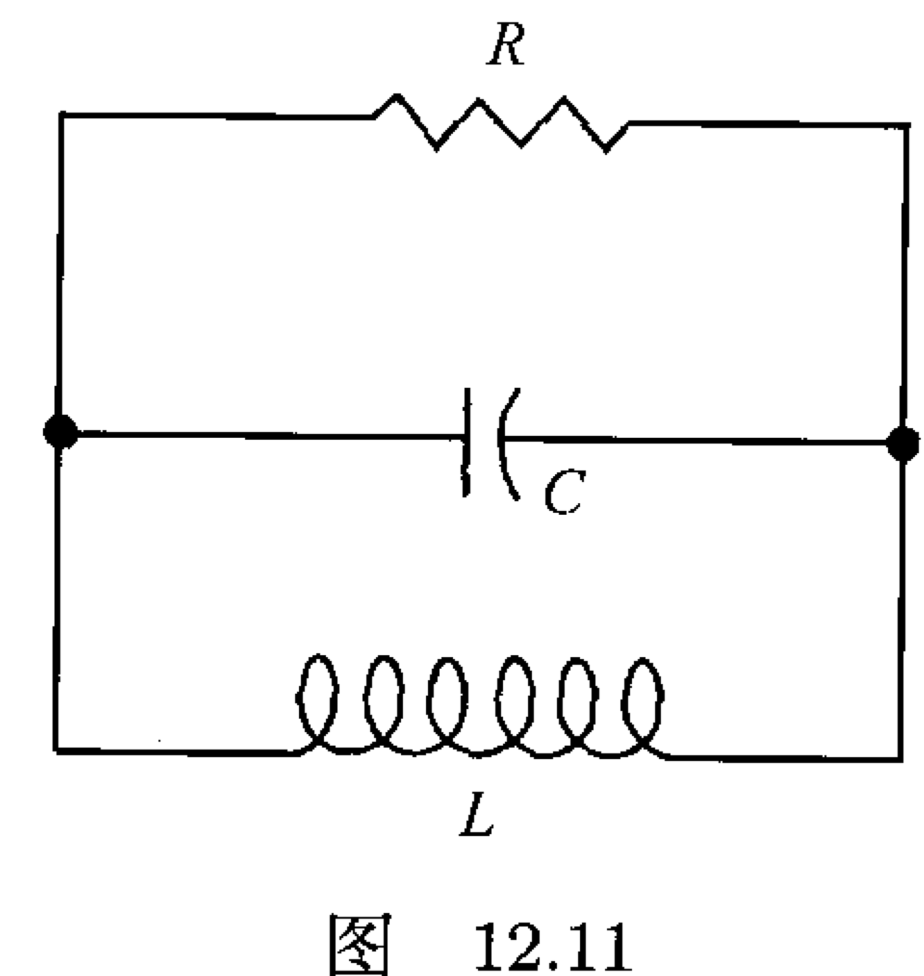


图 12.11

7. 证明：用电线将一个电感和一个电容连成一个闭环的 LC 电路是振荡的.
8. 确定下面的微分方程的相图. 特别地, 要证明只有唯一的非平凡周期解.

$$x'=y-f(x),$$
$$y'=-g(x),$$

这里假设下面所有的条件都满足：

(a) $g(-x)=-g(x)$, 且对所有的 $x\neq 0$ 都有 $xg(x)>0$;

(b) $f(-x)=-f(x)$, 且当 $0<x<a$ 时, 有 $f(x)<0$;

(c) 当 $x>a$ 时, $f(x)$ 是正的且递增;

(d) 当 $x\to\infty$ 时, $f(x)\to\infty$.

9. 考虑系统

$$x'=y,$$
$$y'=a(1-x^4)y-x.$$

(a) 找出所有的平衡点并将它们分类.

(b) 简略地作出相平面.

(c) 描述当 a 变正时出现的分岔.

(d) 证明当 $a>0$ 时, 该系统有唯一的闭轨.

(e) 证明当 $a>0$ 时, 该系统的所有非零解都趋向它的唯一闭轨.

第 13 章　力学中的应用

在本章, 我们将注意力转到微分方程最早期的重要例子, 这些例子事实上与微积分的起源都有关系. 这些方程被牛顿用来推导和统一开普勒三定律. 在本章我们简明推导开普勒定律中的两个, 然后讨论力学中一些更一般的问题.

我们的出发点 —— 牛顿方程, 在整个现代物理史都一直很重要, 而且它们是物理学的**经典力学**部分的根源. 这里给出的具体例子在历史上和科学上都是很重要的. 而且, 我们相当彻底地考虑了一个质点在一个中心引力场中的运动, 这种情形足够简单, 以至于相应的微分方程可以用准确的经典方法 (仅仅是微积分而已!) 显式地解出. 但是考虑到还有许多不能用这种方法解出的复杂力学系统, 我们还对这些问题描述了一个更几何的方法.

13.1　牛顿第二定律

我们将研究一个质点在一个**力场**中的运动. 数学上, $\boldsymbol{F}$ 就是质点的 (构形) 空间中的一个向量场, 构形空间在现在的情形就是 $\mathbb{R}^n$. 从物理观点看, $\boldsymbol{F}(\boldsymbol{X})$ 就是当质点位于 $\boldsymbol{X} \in \mathbb{R}^n$ 时所受到的作用力.

我们主要关心的一个力场例子就是太阳的引力: $\boldsymbol{F}(\boldsymbol{X})$ 就是将位于 $\boldsymbol{X}$ 处的质点吸引到太阳的作用力. 我们将在 13.3 节中深入讨论这个系统.

将物理上的一个力场和数学上的一个微分方程建立联系的就是**牛顿第二定律**: $\boldsymbol{F} = m\boldsymbol{a}$. 这个定律断言一个力场中的质点的运动方式如下: 在任何瞬时, 位于 $\boldsymbol{X}$ 处的质点所受作用力向量 $\boldsymbol{F}$ 等于质点的加速度向量 $\boldsymbol{a}$ 乘上该质点质量 m. 即, 牛顿定律给出了一个二阶微分方程:

$$m\boldsymbol{X}'' = \boldsymbol{F}(\boldsymbol{X}).$$

将这个方程写成一个系统就是

$$\boldsymbol{X}' = \boldsymbol{V},$$

$$\boldsymbol{V}' = \frac{1}{m}\boldsymbol{F}(\boldsymbol{X}),$$

其中 $\boldsymbol{V} = \boldsymbol{V}(t)$ 是该质点的速度. 这是一个 $\mathbb{R}^n \times \mathbb{R}^n$ 上的方程组系统. 这种类型的系统通常称为**具有 n 个自由度的力学系统**.

我们说这个二阶方程的一个解 $\boldsymbol{X}(t) \subset \mathbb{R}^n$ 位于**构形空间**中, 而称这个系统的相应解 $(\boldsymbol{X}(t), \boldsymbol{V}(t)) \in \mathbb{R}^n \times \mathbb{R}^n$ 位于这个系统的**相空间**或**状态空间**中.

例　回忆一下第 2 章中简单的无阻尼调和振子. 在这种情形, 质点在一维空间中运动, 而且它在时刻 t 的位置由一个函数 $x(t)$ 给出, 其中 $x : \mathbb{R} \to \mathbb{R}$. 我们已经看到, 控制质点运动的微分方程就是

$$mx'' = -kx,$$

其中常数 $k > 0$. 即, 力场在点 $x \in \mathbb{R}$ 处由 $-kx$ 给出. ■

例　二维形式的调和振子容许质点在平面上运动, 从而现在确定位置的是一个向量 $\boldsymbol{X}(t) = (x_1(t), x_2(t)) \in \mathbb{R}^2$. 与一维情形一样, 力场仍然是 $\boldsymbol{F}(\boldsymbol{X}) = -k\boldsymbol{X}$, 从而运动方程是一样的

$$m\boldsymbol{X}'' = -k\boldsymbol{X},$$

它在构形空间中的解为

$$\begin{aligned} x_1(t) &= c_1 \cos(\sqrt{k/m}\, t) + c_2 \sin(\sqrt{k/m}\, t), \\ x_2(t) &= c_3 \cos(\sqrt{k/m}\, t) + c_4 \sin(\sqrt{k/m}\, t), \end{aligned}$$

其中 c_j 为常数, 这些解很容易用第 6 章中的方法找出. ■

在讨论牛顿定律的更复杂情况之前, 我们需要回忆多元微积分中的几个概念. 两个向量 $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^n$ 的**点乘**(或**内积**) 记为 $\boldsymbol{X} \cdot \boldsymbol{Y}$, 它的定义是

$$\boldsymbol{X} \cdot \boldsymbol{Y} = \sum_{i=1}^{n} x_i y_i,$$

其中 $\boldsymbol{X} = (x_1, \cdots, x_n)$, $\boldsymbol{Y} = (y_1, \cdots, y_n)$. 于是 $\boldsymbol{X} \cdot \boldsymbol{X} = |\boldsymbol{X}|^2$. 如果 $\boldsymbol{X}, \boldsymbol{Y} : I \to \mathbb{R}^n$ 是光滑函数, 则两个函数相乘的求导法则的另一种形式就是

$$(\boldsymbol{X} \cdot \boldsymbol{Y})' = \boldsymbol{X}' \cdot \boldsymbol{Y} + \boldsymbol{X} \cdot \boldsymbol{Y}',$$

利用坐标函数 x_i, y_i, 可以很容易验证这个等式.

如果 $g : \mathbb{R}^n \to \mathbb{R}$, g 的梯度 $\operatorname{grad} g$ 的定义是

$$\operatorname{grad} g(\boldsymbol{X}) = \left(\frac{\partial g}{\partial x_1}(\boldsymbol{X}), \cdots, \frac{\partial g}{\partial x_n}(\boldsymbol{X}) \right).$$

正如我们在第 9 章中看到的一样, $\operatorname{grad} g$ 是 $\mathbb{R}^n$ 中的一个向量场.

下面, 考虑两个光滑函数 $F : \mathbb{R} \to \mathbb{R}^n$ 和 $g : \mathbb{R}^n \to \mathbb{R}$ 的复合 $g \circ F$. 对 $g \circ F$ 用链式法则可得

$$\frac{\mathrm{d}}{\mathrm{d}t} g(F(t)) = \operatorname{grad} g(F(t)) \cdot F'(t) = \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(F(t)) \frac{\mathrm{d}F_i}{\mathrm{d}t}(t).$$

我们还将用到两个向量 $\boldsymbol{U},\boldsymbol{V}\in\mathbb{R}^3$ 的**叉积**(或**向量积**)$\boldsymbol{U}\times\boldsymbol{V}$. 根据定义

$$\boldsymbol{U}\times\boldsymbol{V}=(u_2v_3-u_3v_2,u_3v_1-u_1v_3,u_1v_2-u_2v_1)\in\mathbb{R}^3.$$

从多元微积分中我们知道

$$\boldsymbol{U}\times\boldsymbol{V}=-\boldsymbol{V}\times\boldsymbol{U}=|\boldsymbol{U}||\boldsymbol{V}|\boldsymbol{N}\sin\theta,$$

其中 $\boldsymbol{N}$ 是同时垂直于 $\boldsymbol{U},\boldsymbol{V}$ 的单位向量, 而且使得向量 $\boldsymbol{U},\boldsymbol{V},\boldsymbol{N}$ 满足 "右手法则". θ 则是 $\boldsymbol{U},\boldsymbol{V}$ 的夹角.

注意, $\boldsymbol{U}\times\boldsymbol{V}=0$ 当且仅当其中一个向量是另一个向量的倍数. 而且, 如果 $\boldsymbol{U}\times\boldsymbol{V}\neq 0$, 则 $\boldsymbol{U}\times\boldsymbol{V}$ 垂直于同时包含 $\boldsymbol{U},\boldsymbol{V}$ 的平面. 如果 $\boldsymbol{U},\boldsymbol{V}$ 都是 $t\in\mathbb{R}$ 的函数, 则两个函数相乘的求导法则的又一种形式就是

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{U}\times\boldsymbol{V})=\boldsymbol{U}'\times\boldsymbol{V}+\boldsymbol{U}\times\boldsymbol{V}',$$

同样可以很容易用分量来验证上式.

13.2 保 守 系 统

物理中出现的很多力场都是通过下面的方式产生的. 存在一个光滑函数 $U:\mathbb{R}^n\to\mathbb{R}$ 使得

$$\boldsymbol{F}(\boldsymbol{X})=-\left(\frac{\partial U}{\partial x_1}(\boldsymbol{X}),\cdots,\frac{\partial U}{\partial x_n}(\boldsymbol{X})\right)=-\mathrm{grad}\,U(\boldsymbol{X}).$$

(其中的负号是一种习惯用法.) 这样的力场称为**保守的**. 此时, 相应的微分方程系统

$$\boldsymbol{X}'=\boldsymbol{V},$$
$$\boldsymbol{V}'=-\frac{1}{m}\mathrm{grad}\,U(\boldsymbol{X})$$

称为一个**保守系统**. 函数 U 称为这个系统的**势能**. [更适当些, U 应该称为**一个**势能, 因为加上任何常数都不会改变力场 $-\mathrm{grad}\,U(\boldsymbol{X})$.]

例 上面的平面调和振子对应于力场 $\boldsymbol{F}(\boldsymbol{X})=-k\boldsymbol{X}$. 这个场是保守的, 它的势能为

$$U(\boldsymbol{X})=\frac{1}{2}k|\boldsymbol{X}|^2.$$

对于任一质量为 m 的运动质点, 它的**动能**定义为

$$K=\frac{1}{2}m|\boldsymbol{V}(t)|^2.$$

注意, 动能依赖于速度, 而势能只是位置的函数. **总能量**(有时就简单地称为**能量**)定义为 $E=K+U$, 它是相空间上的一个函数. 总能量函数之所以在力学上很重要是因为它沿这个系统的任何解曲线都是常数. 或者, 用 9.4 节的语言来说, E 是这个流的一个运动常数, 或一个首次积分. ■

定理　(**能量守恒**) 令 $(\boldsymbol{X}(t),\boldsymbol{V}(t))$ 是保守系统的一条解曲线, 则总能量 E 沿这条解曲线是常数.

证明　为了证明 $E(\boldsymbol{X}(t),\boldsymbol{V}(t))$ 关于 t 是常数, 我们来作计算

$$\begin{aligned}\dot{E} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}m|\boldsymbol{V}(t)|^2+U(\boldsymbol{X}(t))\right)\\ &= m\boldsymbol{V}\cdot\boldsymbol{V}'+(\operatorname{grad}U)\cdot\boldsymbol{X}'\\ &= \boldsymbol{V}\cdot(-\operatorname{grad}U)+(\operatorname{grad}U)\cdot\boldsymbol{V}\\ &= 0.\end{aligned}$$

■

我们还可以将这种类型的系统写成**哈密顿形式**. 从第 9 章可知, $\mathbb{R}^n\times\mathbb{R}^n$ 上的一个系统是一个哈密顿系统, 如果它具有如下的形式：

$$\begin{aligned}x_i' &= \frac{\partial H}{\partial y_i},\\ y_i' &= -\frac{\partial H}{\partial x_i},\end{aligned}$$

其中 $H:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}$ 是系统的**哈密顿函数**. 我们已经看到, 函数 H 沿这样一个系统的解曲线是常数. 为了将保守系统写成哈密顿形式, 我们来作一个简单的变量替换. 我们引入**动量向量**$\boldsymbol{Y}=m\boldsymbol{V}$, 然后令

$$H=K+U=\frac{1}{2m}\sum y_i^2+U(x_1,\cdots,x_n).$$

容易验证, 这就将上面的保守系统化成了哈密顿形式.

13.3　中心力场

一个力场 $\boldsymbol{F}$ 称为**中心的**, 如果对于每个 $\boldsymbol{X}$, $\boldsymbol{F}(\boldsymbol{X})$ 都直接指向原点, 或远离原点. 换句话说, 向量 $\boldsymbol{F}(\boldsymbol{X})$ 总是 X 的一个倍数：

$$\boldsymbol{F}(\boldsymbol{X})=\lambda(\boldsymbol{X})\boldsymbol{X},$$

其中系数 $\lambda(\boldsymbol{X})$ 依赖于 $\boldsymbol{X}$. 我们常常按惯例不考虑质点位于原点的情形; 许多中心力场在原点都没有定义 (或是“无穷”). 我们将在 13.7 节中讨论这些类型的奇

点. 理论上, 随着 $\boldsymbol{X}$ 在球面 $|\boldsymbol{X}|=$ 常数上的取值不同, $\lambda(\boldsymbol{X})$ 也会随之而变化. 但是, 如果力场是保守的, 则情况会不同.

命题 设 $\boldsymbol{F}$ 是一个保守力场. 则下面的几条等价:

(1) $\boldsymbol{F}$ 是中心的;

(2) $\boldsymbol{F}(\boldsymbol{X})=f(|\boldsymbol{X}|)\boldsymbol{X}$;

(3) $\boldsymbol{F}(\boldsymbol{X})=-\operatorname{grad}U(\boldsymbol{X})$, 且 $U(\boldsymbol{X})=g(|\boldsymbol{X}|)$.

证明 假设 (3) 成立. 为了证明 (2), 根据链式法则, 我们有

$$\frac{\partial U}{\partial x_j}=g'(|\boldsymbol{X}|)\frac{\partial}{\partial x_j}\left(x_1^2+x_2^2+x_3^2\right)^{1/2}=\frac{g'(|\boldsymbol{X}|)}{|\boldsymbol{X}|}x_j.$$

令 $f(|\boldsymbol{X}|)=-g'(|\boldsymbol{X}|)/|\boldsymbol{X}|$ 就证明了 (2). 显然 (2) 蕴含 (1). 为了从 (1) 推出 (3), 我们需要证明 U 在每个球面

$$\mathcal{S}_\alpha=\{\boldsymbol{X}\in\mathbb{R}^n||\boldsymbol{X}|=\alpha>0\}$$

上都是常数. 由于 $\mathcal{S}_\alpha$ 上的任何两点都可以用位于 $\mathcal{S}_\alpha$ 中的一条曲线连接, 只需证明 U 在 $\mathcal{S}_\alpha$ 上的每条曲线上都是常数. 于是, 如果 $J\subset\mathbb{R}$ 是一个区间, $\gamma:J\to\mathcal{S}_\alpha$ 是一条光滑曲线, 我们必须证明复合函数 $U\circ\gamma$ 的导数恒等于 0. 在 13.1 节中已经看到, 这个导数就是

$$\frac{\mathrm{d}}{\mathrm{d}t}U(\gamma(t))=\operatorname{grad}U(\gamma(t))\cdot\gamma'(t).$$

由于 F 是中心的, 因而有 $\operatorname{grad}U(\boldsymbol{X})=-F(\boldsymbol{X})=-\lambda(\boldsymbol{X})\boldsymbol{X}$. 于是我们就有

$$\frac{\mathrm{d}}{\mathrm{d}t}U(\gamma(t))=-\lambda(\gamma(t))\gamma(t)\cdot\gamma'(t)=-\frac{\lambda(\gamma(t))}{2}\frac{\mathrm{d}}{\mathrm{d}t}|\gamma(t)|^2=0,$$

其中最后一个等式是因为 $|\gamma(t)|\equiv\alpha$. ■

现在考虑 $\mathbb{R}^3$ 上的一个不一定保守的中心力场. 给定某个时刻 t_0, 记 $\mathcal{P}\subset\mathbb{R}^3$ 为包含位置向量 $\boldsymbol{X}(t_0)$、速度向量 $\boldsymbol{V}(t_0)$ 以及原点的平面 (暂时假设位置向量和速度向量是不共线的). 注意, 作用力向量 $\boldsymbol{F}(\boldsymbol{X}(t_0))$ 也位于 $\mathcal{P}$ 中. 由此似乎可以得出, 质点永远逗留在平面 $\mathcal{P}$ 上. 事实上, 这是成立的:

命题 在 $\mathbb{R}^3$ 的一个中心力场中运动的质点永远逗留在一个包含原点的固定平面上.

证明 假设质点在一个中心力场的作用下的运动路线是 $\boldsymbol{X}(t)$. 我们有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{X}\times\boldsymbol{V})=\boldsymbol{V}\times\boldsymbol{V}+\boldsymbol{X}\times\boldsymbol{V}'=\boldsymbol{X}\times\boldsymbol{X}''=0,$$

最后一个等式是因为 $\boldsymbol{X}''$ 是 $\boldsymbol{X}$ 的倍数. 从而, $\boldsymbol{Y}=\boldsymbol{X}(t)\times\boldsymbol{V}(t)$ 是一个常值向量. 如果 $\boldsymbol{Y}\neq\boldsymbol{0}$, 这就意味着 $\boldsymbol{X},\boldsymbol{V}$ 始终位于垂直于 $\boldsymbol{Y}$ 的平面上, 此时命题得证. 如果 $\boldsymbol{Y}=\boldsymbol{0}$, 则对某个实函数 $g(t)$ 有 $\boldsymbol{X}'(t)=g(t)\boldsymbol{X}(t)$. 这蕴涵着, 与作用在质点上

的力一样, 运动质点的速度向量的方向始终在连接原点和该质点的直线上. 这说明质点始终在过原点的同一条直线上运动. 为了证明这个事实, 令 $\boldsymbol{X}(t)$ 的坐标为 $(x_1(t), x_2(t), x_3(t))$. 于是我们就有 3 个独立的微分方程

$$\frac{\mathrm{d}x_k}{\mathrm{d}t} = g(t)x_k(t), \quad k = 1, 2, 3.$$

积分之后就有

$$x_k(t) = \mathrm{e}^{h(t)}x_k(0), \quad 其中\ h(t) = \int_0^t g(s)\mathrm{d}s.$$

于是, $\boldsymbol{X}(t)$ 始终是 $\boldsymbol{X}(0)$ 的一个倍数, 从而 $\boldsymbol{X}(t)$ 始终在一条固定直线上运动, 当然也就在一个固定平面上运动. ■

向量 $m(\boldsymbol{X} \times \boldsymbol{V})$ 称为系统的**角动量**, 其中 m 为该质点的质量. 根据上一命题的证明, 这个向量也是这个系统的一个守恒量.

推论 (**角动量守恒**) 在一个中心力场中, 角动量沿任何解曲线都是常数. ■

现在我们将注意力集中到一个保守的中心力场. 根据上一命题, 质点始终位于一个平面上, 我们不妨假设这个平面就是 $x_3 = 0$. 此时, 角动量就是向量 $(0, 0, m(x_1v_2 - x_2v_1))$. 令

$$l = m(x_1v_2 - x_2v_1).$$

于是函数 l 沿解为常数. 在平面情形, 我们也称 l 为角动量. 引入极坐标 $x_1 = r\cos\theta, x_2 = r\sin\theta$ 后, 有

$$\begin{aligned} v_1 &= x_1' = r'\cos\theta - r\sin\theta\,\theta', \\ v_2 &= x_2' = r'\sin\theta + r\cos\theta\,\theta'. \end{aligned}$$

于是,

$$\begin{aligned} x_1v_2 - x_2v_1 &= r\cos\theta(r'\sin\theta + r\cos\theta\,\theta') - r\sin\theta(r'\cos\theta - r\sin\theta\,\theta') \\ &= r^2(\cos^2\theta + \sin^2\theta)\theta' \\ &= r^2\theta'. \end{aligned}$$

从而在极坐标中, $l = mr^2\theta'$.

现在我们就来证明开普勒定律中的一个. 记 $A(t)$ 为向量 $\boldsymbol{X}(t)$ 在时间从 t_0 到 t 时所扫过的面积. 在极坐标中, 我们有 $\mathrm{d}A = \frac{1}{2}r^2\mathrm{d}\theta$. 定义**面积速度**为

$$A'(t) = \frac{1}{2}r^2(t)\theta'(t),$$

即位置向量扫过面积的速度. 开普勒观测到行星与太阳的连线在相同时间扫过的面积相等, 我们将此解释成 $A' =$ 常数. 于是, 我们所证明的比它更一般: 对于一个在保守中心力场中运动的任何质点, 这个规律都成立.

现在，对一个中心力场所产生的保守系统，我们就找到了它的两个运动常数，或两个首次积分：总能量和角动量. 在 19 世纪，求解微分方程的方法非常依赖于构造足够多的这种运动常数. 而微分方程经常不存在首次积分，这个事实在 20 世纪逐渐变得显然. 这里罪魁祸首就是混沌；我们将在下面两章中讨论它. 基本上，微分方程的解在一个开集中的混沌行为就排除了在这个集合中存在首次积分的可能.

13.4 牛顿中心力系统

现在我们来专门讨论牛顿中心力系统. 这个系统处理单个行星围绕太阳运动的轨道. 我们假设太阳固定在 $\mathbb{R}^3$ 中的原点，而相对较小的行星在太阳上没有作用力. 太阳在行星上的作用力由**牛顿引力定律**给出，也称为**平方反比率**. 这个定律断言，太阳在位于 $\boldsymbol{X} \in \mathbb{R}^3$ 处的行星上的作用力的大小为 gm_sm_p/r^2，其中 m_s 为太阳质量，m_p 为行星质量，g 为引力常数. 这个力的方向指向太阳. 从牛顿第二定律就可得到微分方程

$$m_p\boldsymbol{X}'' = -gm_sm_p\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3}.$$

为了简明起见，我们变换单位使得所有的常数都为 1，这样微分方程就变得更加简单

$$\boldsymbol{X}'' = \boldsymbol{F}(\boldsymbol{X}) = -\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3},$$

其中 F 代表现在的力场. 写成一个微分方程系统后，我们有

$$\boldsymbol{X}' = \boldsymbol{V},$$

$$\boldsymbol{V}' = -\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3}.$$

这个系统就称为**牛顿中心力系统**. 在本节我们的目标就是描述这个系统的几何；在下节，我们将得出这个系统一个完整的解析解.

显然，这是一个中心力场. 而且，它是保守的，这是因为

$$\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3} = \operatorname{grad} U(\boldsymbol{X}),$$

其中势能 U 为

$$U(\boldsymbol{X}) = -\frac{1}{|\boldsymbol{X}|}.$$

注意，$\boldsymbol{F}(\boldsymbol{X})$ 在 0 处没有定义；事实上，当运动质点与位于原点的静止质点接近碰撞时，这个力场就变成无穷大.

与上节一样, 我们不妨将注意力集中到质点在平面 $\mathbb{R}^2$ 上的运动. 于是, 我们来寻找位于构形空间 $\mathcal{C} = \mathbb{R}^2 - \{0\}$ 中的解. 我们记相空间为 $\mathcal{P} = (\mathbb{R}^2 - \{0\}) \times \mathbb{R}^2$.

我们直观地将相空间看成是 $\mathcal{C}$ 中的每一点 X 处的所有切向量的全体构成的集合. 对一给定的 $\boldsymbol{X} \in \mathbb{R}^2 - \{0\}$, 令 $T_X = \{(\boldsymbol{X}, \boldsymbol{V}) | \boldsymbol{V} \in \mathbb{R}^2\}$. $T_{\boldsymbol{X}}$ 就是构形空间在 $\boldsymbol{X}$ 处的**切平面**. 于是,

$$\mathcal{P} = \bigcup_{\boldsymbol{X} \in \mathcal{C}} T_{\boldsymbol{X}}$$

就是构形空间的**切空间**, 我们可以将它自然地等同成 $\mathbb{R}^4$ 的一个子集.

这个相空间的维数是四. 但是, 我们可以利用两个已知的首次积分 —— 总能量和角动量, 将维数减半. 我们已知能量沿解为常数, 而且

$$E(\boldsymbol{X}, \boldsymbol{V}) = K(\boldsymbol{V}) + U(\boldsymbol{X}) = \frac{1}{2}|\boldsymbol{V}|^2 - \frac{1}{|\boldsymbol{X}|}.$$

记 Σ_h 为 $\mathcal{P}$ 的一个子集, 它由所有满足 $E(\boldsymbol{X}, \boldsymbol{V}) = h$ 的点 $(\boldsymbol{X}, \boldsymbol{V})$ 构成. Σ_h 称为总能量为 h 的一个**能量曲面**. 如果 $h \geqslant 0$, 则 Σ_h 与 $T_{\boldsymbol{X}}$ 相交成一个圆周, 该圆周由满足

$$|\boldsymbol{V}|^2 = 2\left(h + \frac{1}{|\boldsymbol{X}|}\right)$$

的切向量构成. 这些圆周位于 $\boldsymbol{X}$ 处的切平面上, 当 $\boldsymbol{X} \to 0$ 时, 这些圆周的半径趋于 ∞; 而当 $\boldsymbol{X} \to \infty$ 时, 这些圆周的半径下降到 $\sqrt{2h}$.

当 $h < 0$ 时, 能量曲面 Σ_h 的结构则不同. 如果 $|\boldsymbol{X}| > -1/h$, 则在 $T_{\boldsymbol{X}} \cap \Sigma_h$ 中没有向量. 当 $|\boldsymbol{X}| = -1/h$ 时, $T_{\boldsymbol{X}}$ 中只有零向量位于 Σ_h 中. 因而称构形空间中的圆周 $r = -1/h$ 为**零速曲线**. 如果 $\boldsymbol{X}$ 位于零速曲线的内部, 则像前面一样, $T_{\boldsymbol{X}}$ 与能量曲面交成一个由切向量构成的圆周. 图 13.1 给出了 Σ_h 在 $h < 0$ 时的示意图.

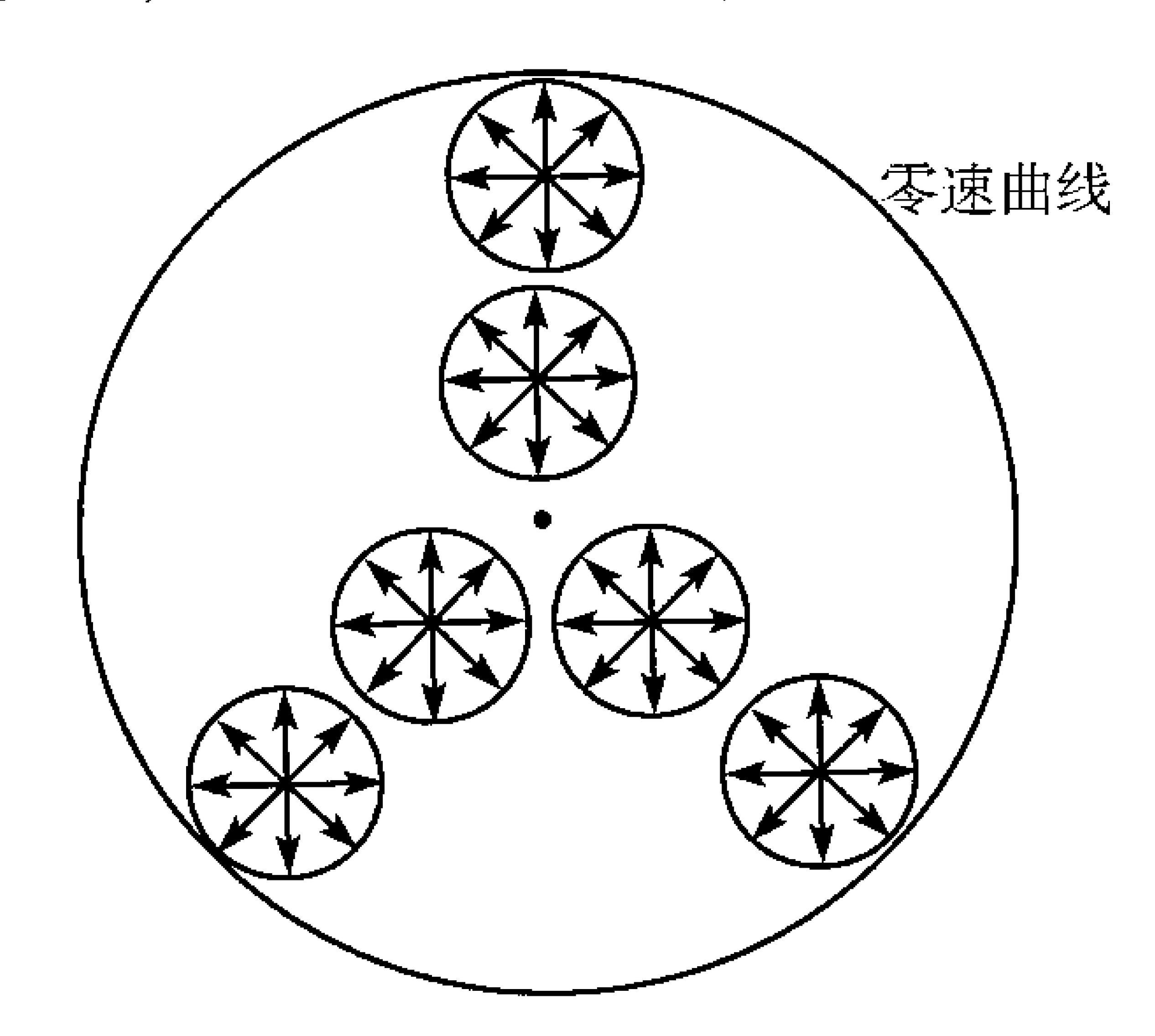


图13.1 在零速曲线内部的每个非零点上, T_X 和能量曲面 Σ_h 交成一个由切向量构成的圆周

现在我们在构形空间中引入极坐标, 而通过如下的方式在切平面上引入新变量 (v_r, v_θ):

$$\boldsymbol{V} = v_r \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + v_\theta \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}.$$

我们有

$$\boldsymbol{V} = \boldsymbol{X}' = r' \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} + r\theta' \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix},$$

于是 $r' = v_r$ 且 $\theta' = v_\theta/r$. 再微分一次, 就有

$$\frac{-1}{r^2}\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = -\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3} = \boldsymbol{V}' = \left(v_r' - \frac{v_\theta^2}{r}\right)\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} + \left(\frac{v_r v_\theta}{r} + v_\theta'\right)\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}.$$

从而在新坐标 (r,θ,v_r,v_θ) 中, 系统变成

$$\begin{aligned} r' &= v_r,\\ \theta' &= v_\theta/r,\\ v_r' &= -\frac{1}{r^2} + \frac{v_\theta^2}{r},\\ v_\theta' &= -\frac{v_r v_\theta}{r}. \end{aligned}$$

在这些坐标下, 总能量为

$$\frac{1}{2}\left(v_r^2 + v_\theta^2\right) - \frac{1}{r} = h,$$

而且角动量由 $l = rv_\theta$ 给出. 记 $\Sigma_{h,l}$ 是相空间中总能量为 h, 角动量为 l 的所有点构成的集合. 为了简单起见, 我们将注意力集中到 $h < 0$ 的情形.

如果 $l = 0$, 必定有 $v_\theta = 0$. 因而, 如果 $\boldsymbol{X}$ 位于零速曲线的内部, 则 $\boldsymbol{X}$ 处的切空间与 $\Sigma_{h,0}$ 正好交于两个向量, 其形式为

$$\pm v_r\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},$$

它们都位于 $\boldsymbol{0}$ 和 $\boldsymbol{X}$ 的连线上, 一个指向 $\boldsymbol{0}$, 另一个远离 $\boldsymbol{0}$. 在零速曲线上, 只有零向量在 $\Sigma_{h,0}$ 中. 于是我们立刻看出 $\Sigma_{h,0}$ 上的每个解都位于一条过原点的直线上. 这个解离开原点后沿一条直线前进直到到达零速曲线, 过了这个时刻, 它就后退回到原点. 事实上, 由于当 $\boldsymbol{X} \to \boldsymbol{0}$ 时, $\Sigma_{h,0}$ 中向量的大小趋于 ∞, 在 (时间向前向后的) 两个方向上, 这些解都在有限时间内到达奇点. 这种类型的解称为**喷出 - 碰撞轨道**.

当 $l \neq 0$ 时, 一种不同的图像就出现了. 给定零速曲线内部的一点 $\boldsymbol{X}$, 有 $v_\theta = l/r$, 于是从总能量的表达式可得,

$$r^2 v_r^2 = 2hr^2 + 2r - l^2. \tag{A}$$

因而方程 (A) 右端关于 r 的二次多项式必须是非负的, 于是这就对那些出现在 $\Sigma_{h,l}$ 中 $\boldsymbol{X}$ 的 r 值加上了限制. 由于 $h < 0$, 这个二次多项式的图像是下凹的. 当

$l^2 > -1/2h$ 时, 它没有实根. 从而此时空间 $\Sigma_{h,l}$ 是空的. 当 $l^2 = -1/2h$ 时, 仅有一个根, 它位于 $r = -1/2h$ 处. 从而这就是此时 $\Sigma_{h,l}$ 中唯一容许的 r 值. 在 (r, θ) 处的切平面中, 有 $v_r = 0, v_\theta = -2hl$, 从而这代表一个圆形闭轨 (如果 $l < 0$ 则顺时针前进, 如果 $l > 0$ 则反时针前进).

如果 $l^2 < -1/2h$, 则这个多项式有一对不同的实根 α, β, 它们满足 $\alpha < -1/2h < \beta$. 注意 $\alpha > 0$. 记 $A_{\alpha,\beta}$ 为构形空间中的环形区域 $\alpha \leqslant r \leqslant \beta$. 于是, 运动在构形空间中就限制在 $A_{\alpha,\beta}$ 上.

命题　假设 $h < 0$ 且 $l^2 < -1/2h$. 则 $\Sigma_{h,l} \subset \mathcal{P}$ 是一个二维环面.

证明　对于每个 $\boldsymbol{X} \in A_{\alpha,\beta}$, 我们来计算位于 $T_{\boldsymbol{X}} \cap \Sigma_{h,l}$ 中切向量的集合. 如果 $\boldsymbol{X}$ 位于环带的边界上, 方程 (A) 右端的二次项为零, 从而 $v_r = 0, v_\theta = l/r$. 于是, 当 $\boldsymbol{X}$ 位于环带的边界上时, $T_{\boldsymbol{X}} \cap \Sigma_{h,l}$ 中有唯一的切向量. 当 $\boldsymbol{X}$ 在 $A_{\alpha,\beta}$ 的内部时, 有

$$v_r^{\pm} = \pm\frac{1}{r}\sqrt{2hr^2 + 2r - l^2},\ v_\theta = l/r,$$

从而在这种情形, $T_{\boldsymbol{X}} \cap \Sigma_{h,l}$ 含有一对向量. 注意, 由于对所有的 $\boldsymbol{X}$, v_θ 具有相同的符号, 这些向量在 $A_{\alpha,\beta}$ 中要么全是顺时针的, 要么全是反时针的 (见图 13.2). 从而, 我们可以认为 $\Sigma_{h,l}$ 由 $A_{\alpha,\beta}$ 上的一对图像给出：一个由 v_r^+ 给出的正图像以及一个由 v_r^- 给出的负图像, 而且它们沿两个边界圆周 $r = \alpha$ 和 $r = \beta$ 连接成一体. (当然, "真实的" 图形是 $\mathbb{R}^4$ 的一个子集.) 这就得到了一个环面. ■

我们会忍不住认为环面上由 $r = \alpha$ 和 $r = \beta$ 给出的两条曲线就是该系统的两条闭轨, 然而事实并非如此. 这是因为, 当 $r = \alpha$ 时, 我们有

$$v_r' = -\frac{1}{\alpha^2} + \frac{v_\theta^2}{\alpha} = \frac{1}{\alpha^3}(-\alpha + l^2).$$

但是, 由于方程 (A) 的右端在 $r = \alpha$ 上为零, 我们有

$$2h\alpha^2 + 2\alpha - l^2 = 0,$$

从而,

$$-\alpha + l^2 = (2h\alpha + 1)\alpha.$$

因为 $\alpha < -1/2h$, 于是就有, 当 $r = \alpha$ 时, $r'' = v_r' > 0$, 因而 $\Sigma_{h,l}$ 中的解曲线在与 $r = \alpha$ 相交时, 其 r 坐标达到一个极小值. 类似地, 沿 $r = \beta$, 解曲线的 r 坐标达到一个极大值.

$A_{\alpha,\beta}$ 中解的行为必定是如图 13.3 所示. 注意, 可以容易地证明, 这些解曲线关于原点旋转之后仍然是解曲线, 因而所有解的行为都具有对称性.

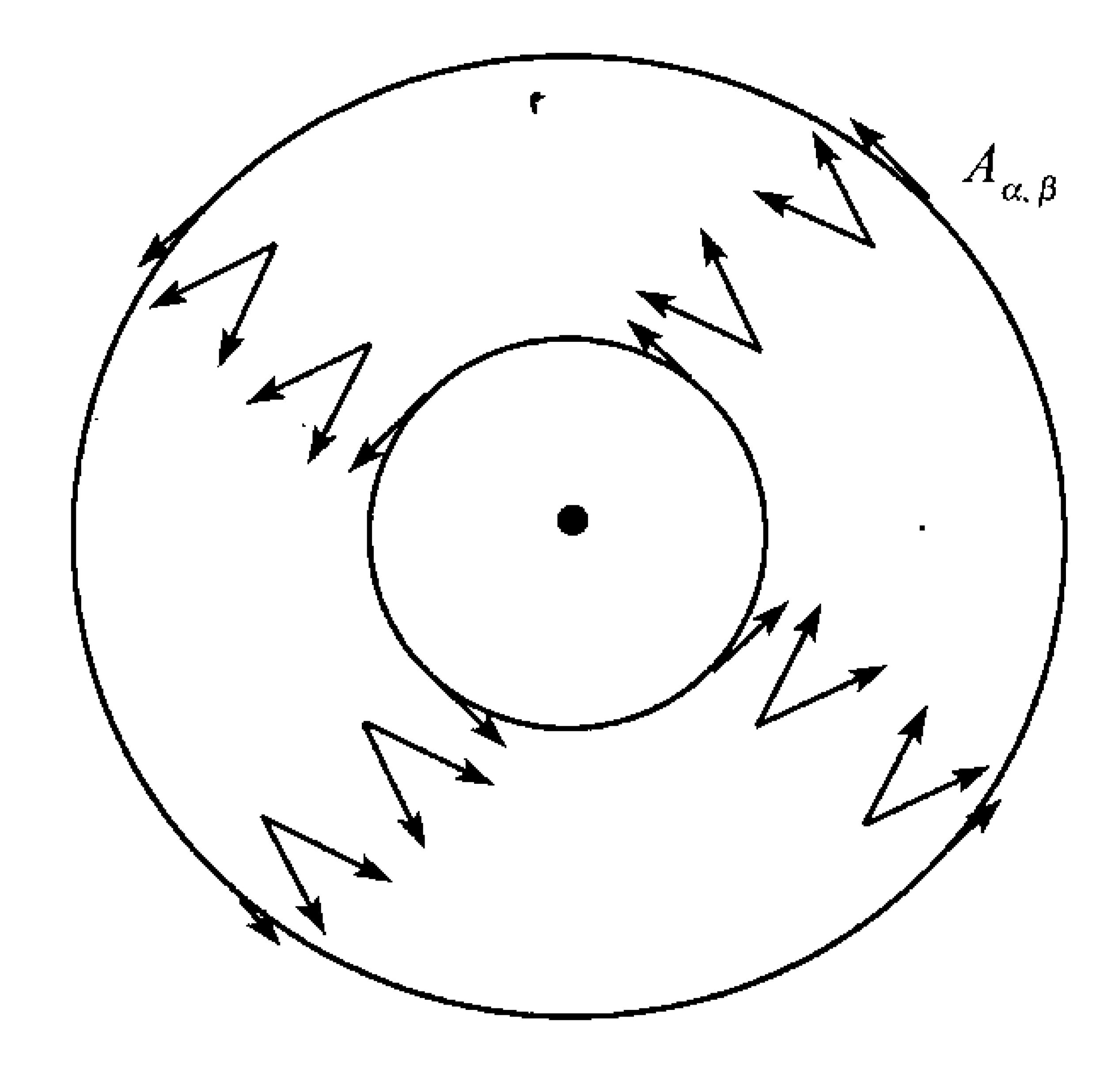


图13.2 $\Sigma_{h,l}$ 中的一些向量

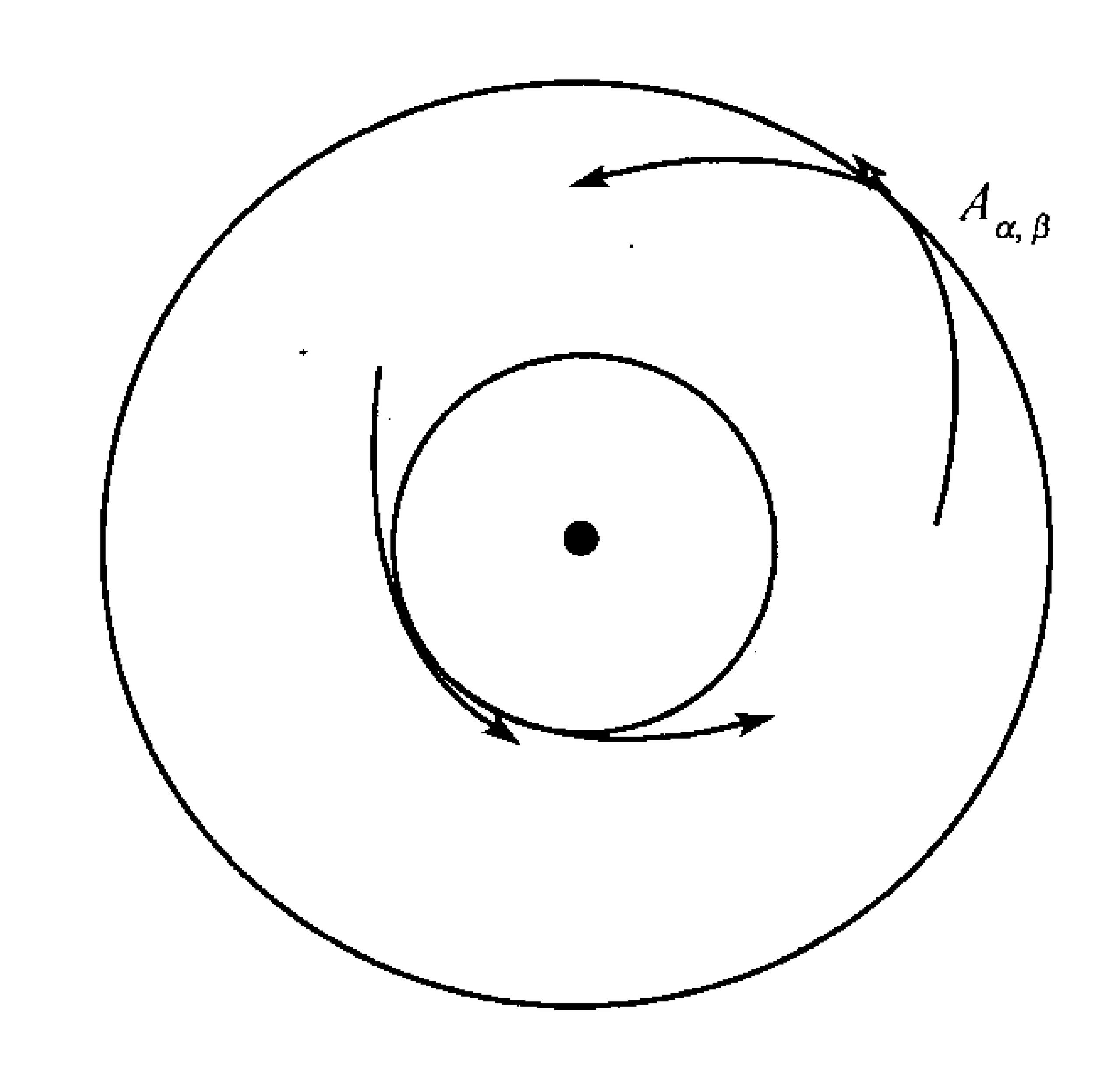


图13.3 $\Sigma_{h,l}$ 中与$r=\alpha$ 或$r=\beta$ 相交的解

但是，还可以说得更多些. 事实上，这些解的每一个都位于一条闭轨上，而且它在构形空间中描绘出一个椭圆. 为了看出这一点，我们需要转到分析方面.

13.5 开普勒第一定律

对多数的非线性力学系统，上节中所作的几何式分析就是我们的全部期望. 但是，在牛顿中心力系统中，我们很幸运：几个世纪前人们就已经知道，可以显式地写出这个系统的解.

考虑这个微分方程的一个特定的解曲线. 该系统有两个运动常数，即，角动量 l 和总能量 E. $l=0$ 的情形就得到喷出-碰撞解，正如我们在上节所见. 因而我们假设 $l \neq 0$. 我们将证明，在构形空间的极坐标中，一个具有非零角动量的解在一条由 $r(1+\epsilon\cos\theta)=\kappa$ 所给出的曲线上，其中 ϵ,κ 为常数. 将这个方程在笛卡儿坐标中重新写出就可以看出，这个方程定义了一条圆锥曲线. 这一事实就是著名的**开普勒第一定律**.

为了证明这个事实，注意，$r^2\theta'$ 是常值而且非零. 所以 θ' 的符号沿每条解曲线保持不变. 从而，θ 随时间要么总是增加，要么总是减少. 于是，沿解曲线，我们可以将 r 看成 θ 的一个函数.

令 $W(t)=1/r(t)$; 则 W 也是 θ 的一个函数. 注意，$W=-U$. 下面的命题给出了动能的一个有用的表达式.

命题 系统的动能可由下式给出：

$$K=\frac{l^2}{2}\left(\left(\frac{\mathrm{d}W}{\mathrm{d}\theta}\right)^2+W^2\right).$$

证明 在极坐标中, 我们有

$$K=\frac{1}{2}\left((r')^2+(r\theta')^2\right).$$

由于 $r=1/W$, 我们还有

$$r'=\frac{-1}{W^2}\frac{\mathrm{d}W}{\mathrm{d}\theta}\theta'=-l\frac{\mathrm{d}W}{\mathrm{d}\theta}.$$

最后还有

$$r\theta'=\frac{l}{r}=lW.$$

将它们代入 K 的表达式就可完成证明. ■

现在, 沿系统的解曲线, 我们找出一个联系 W 和 θ 的微分方程. 由 $K=E-U=E+W$, 再根据上一命题可得

$$\left(\frac{\mathrm{d}W}{\mathrm{d}\theta}\right)^2+W^2=\frac{2}{l^2}(E+W). \tag{B}$$

在上式两边关于 θ 求导数后, 再除以 $2\mathrm{d}W/\mathrm{d}\theta$, 最后再利用 $\mathrm{d}E/\mathrm{d}\theta=0$(能量守恒), 就得到

$$\frac{\mathrm{d}^2W}{\mathrm{d}\theta^2}+W=\frac{1}{l^2},$$

其中 $1/l^2$ 是一个常数.

注意, 这个方程就是具有常值外力 $1/l^2$ 的调和振子方程. 从初等微积分中可知, 这个二阶方程的解可以写成下面的形式:

$$W(\theta)=\frac{1}{l^2}+A\cos\theta+B\sin\theta,$$

或者, 等价地写成

$$W(\theta)=\frac{1}{l^2}+C\cos(\theta+\theta_0),$$

其中常数 C 和 θ_0 与 A 和 B 有关.

如果将这个表达式代入方程 (B), 然后解出 C(比如, 在 $\theta+\theta_0=\pi/2$ 处) 可得

$$C=\pm\frac{1}{l^2}\sqrt{1+2l^2E}.$$

将这个式子代入上面的解, 就有

$$W(\theta)=\frac{1}{l^2}\left(1\pm\sqrt{1+2l^2E}\cos(\theta+\theta_0)\right).$$

没有必要考虑根号前面的两个符号, 因为我们有

$$\cos(\theta+\theta_0+\pi)=-\cos(\theta+\theta_0).$$

而且, 如果将变量 θ 替换成 $\theta-\theta_0$, 则我们可以将任何特定的解都写成如下形式:

$$\frac{1}{l^2}\left(1+\sqrt{1+2l^2E}\cos\theta\right).$$

这看起来相当复杂. 但是, 从解析几何中 (或从本章的习题 2) 可知, 极坐标中一条圆锥曲线的方程为

$$\frac{1}{r}=\frac{1}{\kappa}(1+\epsilon\cos\theta).$$

这里 κ 是**正焦弦**, 而 $\epsilon\geqslant 0$ 为该圆锥曲线的**离心率**. 原点是一个焦点, 三种情形 $\epsilon>1,\epsilon=1,\epsilon<1$ 分别对应于双曲线、抛物线、椭圆. $\epsilon=0$ 的情形是圆周. 在我们的情形, 有

$$\epsilon=\sqrt{1+2l^2E},$$

于是这三种不同的情形分别在 $E>0,E=0,E<0$ 时出现. 我们已经证明:

定理 (**开普勒第一定律**) 在牛顿引力定律作用下的质点运动轨迹是一条圆锥曲线, 其离心率为

$$\sqrt{1+2l^2E}.$$

根据 $E>0,E=0$ 或 $E<0$ 的不同情形, 这些轨迹分别位于双曲线、抛物线或椭圆上. ■

13.6 二体问题

我们现在简要地讨论一下一个最初看起来更困难的问题 ——**二体问题**. 在这个系统中, 我们有两个质点, 它们在空间的运动由相互之间的万有吸引力所决定.

记 $\boldsymbol{X}_1,\boldsymbol{X}_2$ 为 $\mathbb{R}^3$ 中两个质量为 m_1,m_2 的质点的位置. 于是, $\boldsymbol{X}_1=(x_1^1,x_2^1,x_3^1)$, $\boldsymbol{X}_2=(x_1^2,x_2^2,x_3^2)$. 根据牛顿引力定律, 我们可找出它们的运动方程为

$$m_1\boldsymbol{X}_1''=gm_1m_2\frac{\boldsymbol{X}_2-\boldsymbol{X}_1}{|\boldsymbol{X}_2-\boldsymbol{X}_1|^3},$$

$$m_2\boldsymbol{X}_2''=gm_1m_2\frac{\boldsymbol{X}_1-\boldsymbol{X}_2}{|\boldsymbol{X}_1-\boldsymbol{X}_2|^3}.$$

我们从一个生活在第一个质点的观察者的角度来考察这些方程. 令 $\boldsymbol{X}=\boldsymbol{X}_2-\boldsymbol{X}_1$. 于是就有

$$\begin{aligned}\boldsymbol{X}'' &= \boldsymbol{X}_2'' - \boldsymbol{X}_1'' \\ &= gm_1\frac{\boldsymbol{X}_1 - \boldsymbol{X}_2}{|\boldsymbol{X}_1 - \boldsymbol{X}_2|^3} - gm_2\frac{\boldsymbol{X}_2 - \boldsymbol{X}_1}{|\boldsymbol{X}_2 - \boldsymbol{X}_1|^3} \\ &= -g(m_1 + m_2)\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3}.\end{aligned}$$

但是, 这正是牛顿中心力问题, 只不过所选的常数不同而已.

因而, 为了求解二体问题, 我们首先确定这个中心力问题的解 $\boldsymbol{X}(t)$. 而这就使得 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 的微分方程的右端都是 t 的已知函数, 然后再简单地积分两次就可找到 $\boldsymbol{X}_1(t)$ 和 $\boldsymbol{X}_2(t)$.

下面是另一种将二体问题约化成牛顿中心力系统的方法. 二体系统的**质心**为向量

$$\boldsymbol{X}_c = \frac{m_1\boldsymbol{X}_1 + m_2\boldsymbol{X}_2}{m_1 + m_2}.$$

计算表明 $\boldsymbol{X}_c'' = \boldsymbol{0}$. 于是我们必定有 $\boldsymbol{X}_c = \boldsymbol{A}t + \boldsymbol{B}$, 其中 $\boldsymbol{A}, \boldsymbol{B}$ 是 $\mathbb{R}^3$ 中的两个固定向量. 这说明系统的质心作匀速直线运动.

现在我们作坐标变换使得系统的原点位于 $\boldsymbol{X}_c$. 即, 令 $\boldsymbol{Y}_j = \boldsymbol{X}_j - \boldsymbol{X}_c, j = 1, 2$. 从而对所有的 t, 都有 $m_1\boldsymbol{Y}_1(t) + m_2\boldsymbol{Y}_2(t) = \boldsymbol{0}$. 将微分方程用 $\boldsymbol{Y}_j$ 重新写出就可得到

$$\boldsymbol{Y}_1'' = -\frac{gm_2^3}{(m_1 + m_2)^3}\frac{\boldsymbol{Y}_1}{|\boldsymbol{Y}_1|^3},$$

$$\boldsymbol{Y}_2'' = -\frac{gm_1^3}{(m_1 + m_2)^3}\frac{\boldsymbol{Y}_2}{|\boldsymbol{Y}_2|^3},$$

这就得到了一对中心力问题. 但是由于有 $m_1\boldsymbol{Y}_1(t) + m_2\boldsymbol{Y}_2(t) = \boldsymbol{0}$, 我们只需求解其中的一个即可.

13.7 吹胀奇点

在牛顿中心力问题中位于原点的奇点是我们第一次遇到这种情形. 通常我们的向量场在整个 $\mathbb{R}^n$ 上都是定义好的. 在力学上, 有时可以通过适当地变换变量和时间标度而将这种奇点消除. 在牛顿中心力系统中, 这可以利用 McGehee 引入的一种变量替换来达到 [32].

我们首先调整变量

$$u_r = r^{1/2}v_r,$$

$$u_\theta = r^{1/2} v_\theta.$$

在这些变量下，系统变成

$$\begin{aligned}
r' &= r^{-1/2} u_r, \\
\theta' &= r^{-3/2} u_\theta, \\
u_r' &= r^{-3/2} \left(\frac{1}{2} u_r^2 + u_\theta^2 - 1 \right), \\
u_\theta' &= r^{-3/2} \left(-\frac{1}{2} u_r u_\theta \right).
\end{aligned}$$

此时，原点仍然是一个奇点，但是要注意，后面三个方程都是乘上了 $r^{-3/2}$. 简单地将向量场乘上 $r^{3/2}$, 我们就可以将这些项消除. 这样做了之后，系统的解曲线依然是一样的，只不过参数发生了变化.

准确地说，我们通过如下的规则引入一个新的时间变量 τ:

$$\frac{\mathrm{d}t}{\mathrm{d}\tau} = r^{3/2}.$$

根据链式法则，有

$$\frac{\mathrm{d}r}{\mathrm{d}\tau} = \frac{\mathrm{d}r}{\mathrm{d}t} \frac{\mathrm{d}t}{\mathrm{d}\tau},$$

对其他变量也有类似的关系. 在这个新时间标度下，系统变成

$$\begin{aligned}
\dot{r} &= r u_r, \\
\dot{\theta} &= u_\theta, \\
\dot{u}_r &= \frac{1}{2} u_r^2 + u_\theta^2 - 1, \\
\dot{u}_\theta &= -\frac{1}{2} u_r u_\theta,
\end{aligned}$$

其中变量上面的点代表关于 τ 求导数. 注意，当 r 很小时，$\mathrm{d}t/\mathrm{d}\tau$ 接近零，从而在原点附近时间 t 要比 "时间" τ 慢得多.

原点不再是这个系统的奇点. 我们已经将奇点 "吹胀" (blow up) 起来，变成了一个新集合 $r = 0$, 此时 θ, u_r, u_θ 是任意的. 在这个集合上，系统是完全定义好的. 事实上，由于当 $r = 0$ 时，$r' = 0$, 集合 $r = 0$ 是这个流的一个不变集. 于是，我们在 $r = 0$ 上引入了一个虚构流. 虽然在 $r = 0$ 上的解在真正的系统中没有意义，但是，根据解的连续性，从中可以得出关于奇点附近解的行为的许多信息.

我们不必考虑所有的 $r=0$, 因为在新坐标中总能量关系变成

$$h_r = \frac{1}{2}\left(u_r^2 + u_\theta^2\right) - 1.$$

在集合 $r=0$ 上, 只有由

$$u_r^2 + u_\theta^2 = 2, \quad \theta \text{ 是任意的}$$

定义的集合 Λ 是有关系的. 集合 Λ 称为这个系统的**碰撞曲面**; 因为在新坐标中, 任何逼近 $r=0$ 的解都要逼近 Λ, 因而解在 Λ 上的行为就反应了解在奇点附近是如何运动的. 注意, Λ 是一个二维环面：它由 θ 方向的一个圆周与 $u_r u_\theta$ 平面的一个圆周构成.

在 Λ 上, 这个系统约化成

$$\begin{aligned}
\dot{\theta} &= u_\theta, \\
\dot{u}_r &= \frac{1}{2}u_\theta^2, \\
\dot{u}_\theta &= -\frac{1}{2}u_r u_\theta,
\end{aligned}$$

其中我们利用能量关系简化了 $\dot{u}_r$. 这个系统很容易分析. 如果 $u_\theta \neq 0$, 就有 $\dot{u}_r > 0$. 在 Λ 中, 沿任何 $u_\theta \neq 0$ 的解, u_r 坐标都必定增加.

另一方面, $u_\theta = 0$ 都对应于系统的平衡点. 系统在两个圆周上都是平衡点, 其中一个圆周由 $u_\theta = 0, u_r = \sqrt{2}$ 给出 (θ 任意), 另外一个圆周由 $u_\theta = 0, u_r = -\sqrt{2}$ 给出 (θ 任意). 记 $C^\pm$ 为这两个圆周, 其中在 $C^\pm$ 上, $u_r = \pm\sqrt{2}$. 由于 u_r 沿解增加, 因而所有解都是从 C^- 前进到 C^+.

为了完全了解 Λ 上的流, 我们在每个 $u_r u_\theta$ 平面上引入角变量 ψ 如下：

$$\begin{aligned}
u_r &= \sqrt{2}\sin\psi, \\
u_\theta &= \sqrt{2}\cos\psi.
\end{aligned}$$

现在, 这个环面由 θ 和 ψ 所参数化. 在 $\theta\psi$ 坐标中, 系统变成

$$\begin{aligned}
\dot{\theta} &= \sqrt{2}\cos\psi, \\
\dot{\psi} &= \frac{1}{\sqrt{2}}\cos\psi.
\end{aligned}$$

现在圆周 $C^\pm$ 由 $\psi = \pm\pi/2$ 给出. 在这些方程中消去时间后就得到

$$\frac{d\psi}{d\theta} = \frac{1}{2}.$$

因而, 在 $\theta\psi$ 坐标中观察时, 所有非平衡解都有常值斜率 1/2(见图 13.4).

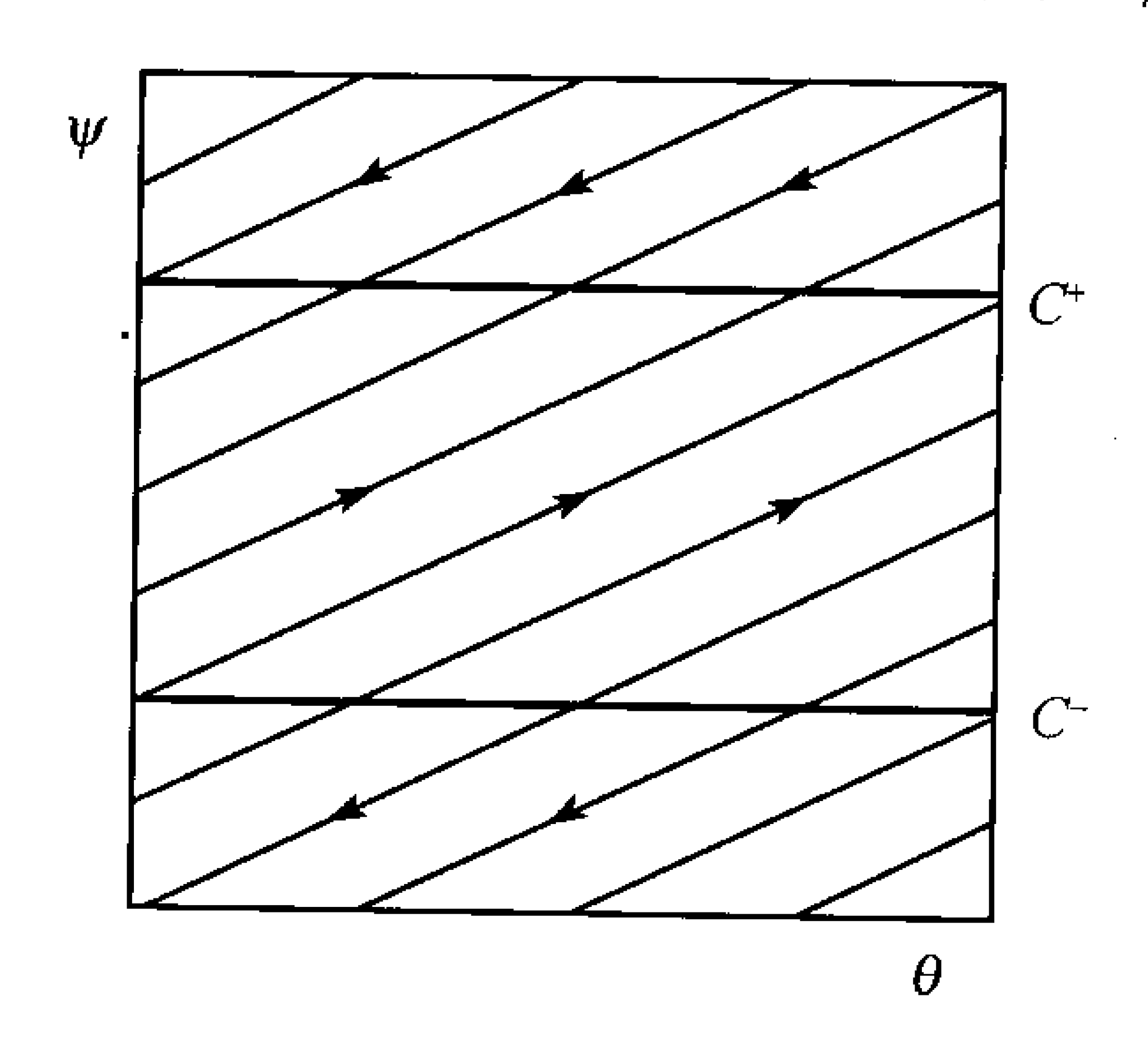


图 13.4　在 $\theta\psi$ 坐标中 Λ 上的解. 注意 θ, ψ 都是对 mod2π 定义的, 因而这个正方形的对边被等同后形成一个环面

现在回忆一下 13.4 节中描述的喷出–碰撞解. 每个这种解沿构形空间中的射线 $\theta = \theta^*$ 先离开原点然后再返回. 这个解离开时 $v_r > 0$(从而 $u_r > 0$), 而返回时 $v_r < 0$ (从而 $u_r < 0$). 于是, 在新的四维坐标系中, 这个解同时是平衡点 $(0, \theta^*, \sqrt{2}, 0)$ 的一条不稳定曲线和平衡点 $(0, \theta^*, -\sqrt{2}, 0)$ 的一条稳定曲线 (见图 13.5). 对附近的非碰撞解又会发生什么呢? 它们先逼近 "较低的" 平衡点 (其 $\theta = \theta^*, u_r = -\sqrt{2}$.), 然后沿不稳定曲线中的一支穿过这点上升到 "较高的" 平衡点 (其 $\theta = \theta^*, u_r = \sqrt{2}$.), 最后沿这个平衡点的不稳定曲线附近离开. 如果在构形空间中解释这个运动, 我们就看到, 每个接近碰撞的解先接近原点, 然后在 θ 增加或减少 2π 个单位后再逃离. 当然, 我们已经知道这一点, 这是因为这些解在一个很扁的椭圆上绕原点飞速旋转.

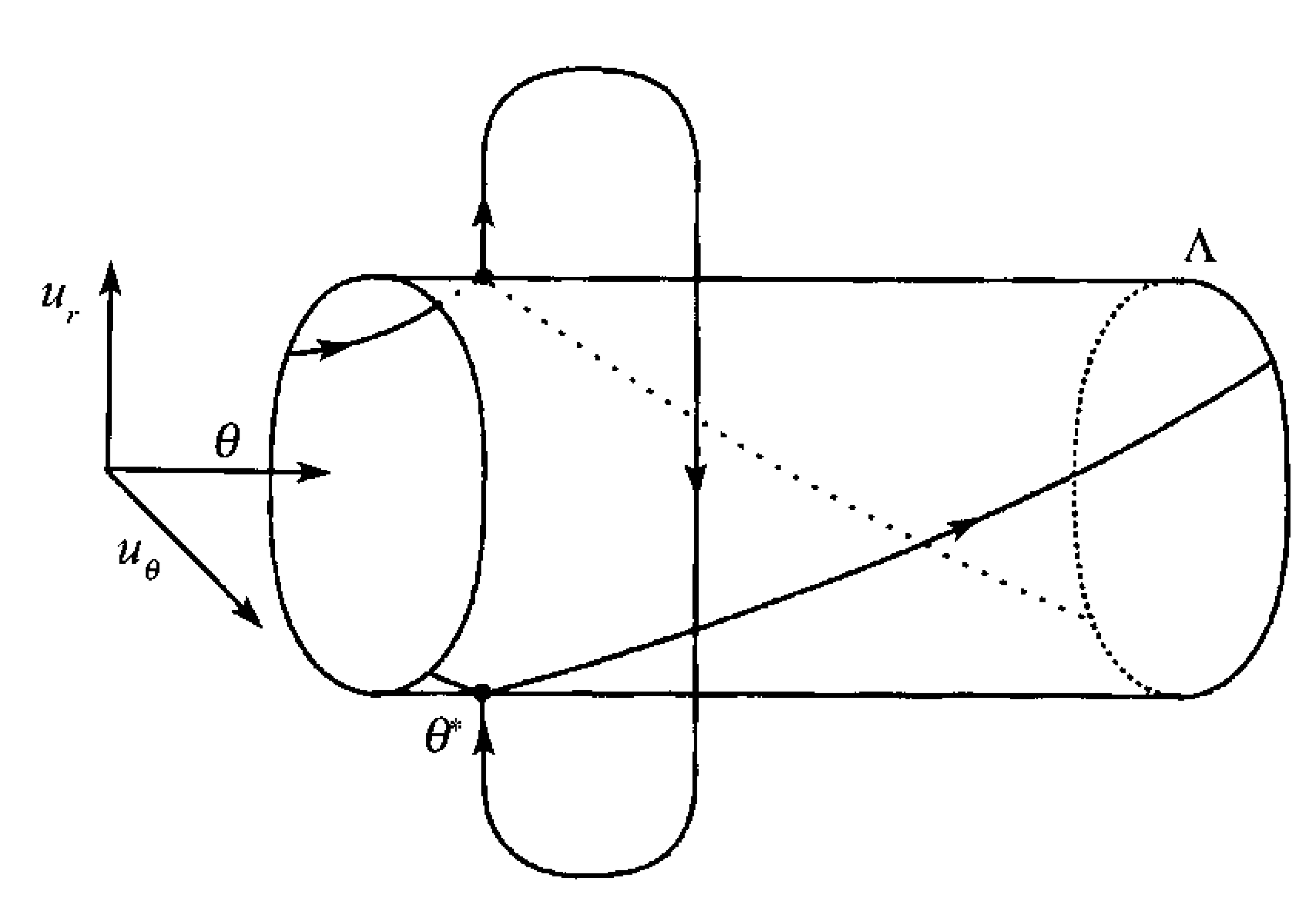


图 13.5　在区域 $r > 0$ 离开后返回到 Λ 的喷出–碰撞解和碰撞曲面上的连结轨道

13.8 探索: 其他中心力问题

在这个探索中, 我们考虑 (非牛顿) 中心力问题, 此时的势能为

$$U(\boldsymbol{X}) = -\frac{1}{|\boldsymbol{X}|^\nu},$$

其中 $\nu > 1$. 主要目的就是了解近碰撞解.

(1) 在极坐标 $(r, \theta, v_r, v_\theta)$ 中写出这个系统. 明确地陈述总能量和角动量的表达式.

(2) 利用计算机, 在 $h < 0$ 和 $l \neq 0$ 时, 研究这个系统的解的行为.

(3) 通过下面的变量替换和时间标度的适当改变, 吹胀原点处的奇点,

$$u_r = r^{\nu/2} v_r, \quad u_\theta = r^{\nu/2} v_\theta,$$

并写出新系统.

(4) 计算碰撞曲面 Λ 上的向量场, 其中 Λ 是根据总能量关系, 由 $r = 0$ 确定.

(5) 描述在 Λ 上当 $\nu = 2$ 时出现的分岔.

(6) 对所有的 $\nu > 1$, 描述 $\Sigma_{h,l}$ 的结构.

(7) 描述当 ν 通过 2 时, $\Sigma_{h,l}$ 结构的变化.

(8) 当 $\nu > 2$ 时, 描述 $\Sigma_{h,l}$ 中解的行为.

(9) 假设 $1 < \nu < 2$. 描述原点处奇点附近的解的行为.

(10) 利用这个系统的解关于原点的旋转保持的性质, 在 $h < 0$ 且 $l \neq 0$ 时, 描述 $\Sigma_{h,l}$ 中解的行为.

13.9 探索: 量子力学系统的经典极限

在这个探索中我们研究各向异性的开普勒问题. 这是依赖于一个参数 μ 的两个自由度的经典力学系统. 当 $\mu = 1$ 时, 这个系统约化成 13.4 节中讨论的牛顿中心力系统. 当 $\mu > 1$ 时, 该系统出现了一些各向异性, 因而它不再是一个中心力场. 与中心力系统一样, 我们仍然有一些喷出 - 碰撞轨道, 但附近轨道的行为与 $\mu = 1$ 的情形十分不同.

各向异性的开普勒问题首先是由 Gutzwiller 作为某种量子力学系统的经典力学逼近而引入的. 特别地, 当人们寻找电子在一个半导体的施主杂质附近的束缚态时, 这种系统就自然地产生了. 这里, 势能由通常的库仑场决定, 然而由于固体中的电子能带结构, 动能成为各向异性的. 等价地, 我们可以认为这个系统具有一个各向异性的势能函数. Gutzwiller 认为这种情形类似于电子在一个方向的质量要比其他方向来得大. 对于这个工作在量子力学中应用的更多背景, 参见 [22].

各向异性的开普勒系统为

$$x'' = \frac{-\mu x}{(\mu x^2 + y^2)^{3/2}},$$
$$y'' = \frac{-y}{(\mu x^2 + y^2)^{3/2}},$$

其中 μ 是一个参数, 我们假设 $\mu > 1$.

(1) 证明这个系统是保守的, 其势能为

$$U(x,y) = \frac{-1}{\sqrt{\mu x^2 + y^2}},$$

并写出总能量的明确表达式.

(2) 当能量 $h < 0$ 时, 描述能量曲面 Σ_h 的几何结构.

(3) 限制在负能量的情形, 证明与零速曲线相交的解只有位于构形空间的 x 轴和 y 轴上的 (直线) 喷出–碰撞轨道.

(4) 证明角动量不再是这个系统的一个首次积分.

(5) 在极坐标中重新写出这个系统.

(6) 像在牛顿中心力问题 (13.7 节) 中一样, 通过变换变量和时间标度吹胀奇点, 并写出一个在 $r = 0$ 处没有任何奇点的新系统.

(7) 描述碰撞曲面 Λ(Σ_h 与新坐标中的 $r = 0$ 的交集) 的结构. 特别地, 为什么有人称这个曲面为葫芦环?

(8) 找出 Λ 上的所有平衡点并确定这些点处的线性化系统的特征值. 确定 Λ 上的哪些平衡点是汇点、源点和鞍点.

(9) 解释当 $\mu = 9/8$ 时出现的分岔.

(10) 找出 Λ 上一个沿所有非平衡解都非减的函数.

(11) 确定碰撞曲面上鞍点的稳定和不稳定曲线的命运. 提示：在该曲面上重新写出方程, 消去时间参数, 当解在 Λ 上攀爬时, 估计它们的斜率.

(12) 当 $\mu > 9/8$ 时, 用定性语言描述 x 轴上喷出–碰撞轨道附近的解接近碰撞时的行为. 特别地, 在构形空间中, 它们是如何逃离原点的? 当解在 y 轴附近行进时, 它们是如何逼近碰撞的?

习　题

1. 下面哪些 $\mathbb{R}^2$ 上的力场是保守的?

(a) $F(x,y) = (-x^2, -2y^2)$;

(b) $F(x,y) = (x^2 - y^2, 2xy)$;

(c) $F(x,y)=(x,0)$.

2. 证明当 $\epsilon>1, \epsilon=1$ 和 $\epsilon<1$ 时, 方程

$$\frac{1}{r}=\frac{1}{h}(1+\epsilon\cos\theta)$$

分别确定一条双曲线、抛物线和椭圆.

3. 在牛顿中心力系统中, 考虑一个质点在时刻 $t=0$ 从原点直接离开的情形. 找出这个解的一个明确表达式, 并讨论该质点的相应运动. 对哪些初值条件, 质点会最终返回?

4. 在牛顿中心力系统中, 当 $h>0$ 和 $h=0$ 时, 描述 $\Sigma_{h,l}$ 的几何结构.

5. 设 $\boldsymbol{F}(\boldsymbol{X})$ 为 $\mathbb{R}^3$ 中的一个力场. 令 $\boldsymbol{X}_0, \boldsymbol{X}_1$ 为 $\mathbb{R}^3$ 中的两个点, $\boldsymbol{Y}(s), s_0 \leqslant s \leqslant s_1$ 为 $\mathbb{R}^3$ 中一条从 $\boldsymbol{X}_0$ 到 $\boldsymbol{X}_1$ 的路径, 其中的 s 为弧长参数. 将质点沿这条路径移动所做的**功**定义为积分

$$\int_{s_0}^{s_1} \boldsymbol{F}(\boldsymbol{Y}(s))\cdot\boldsymbol{Y}'(s)\mathrm{d}s,$$

其中 $\boldsymbol{Y}'(s)$ 为这条路径的 (单位) 切向量. 证明力场为守恒的当且仅当做功与路径无关. 事实上, 如果 $\boldsymbol{F}=-\mathrm{grad}\,\boldsymbol{V}$, 则做的功为 $\boldsymbol{V}(\boldsymbol{X}_1)-\boldsymbol{V}(\boldsymbol{X}_0)$.

6. 描述由

$$\boldsymbol{X}''=-\frac{\boldsymbol{X}}{|\boldsymbol{X}|^4}$$

给出的非牛顿中心力系统的解.

7. 讨论方程

$$\boldsymbol{X}''=\frac{\boldsymbol{X}}{|\boldsymbol{X}|^3}$$

的解. 这个方程对应于原点处的一个排斥力, 而不是吸引力.

8. 这个以及下面两个问题都是讨论二体问题. 令势能为

$$U=\frac{gm_1m_2}{|\boldsymbol{X}_2-\boldsymbol{X}_1|},$$

而且

$$\mathrm{grad}_j(U)=\left(\frac{\partial U}{\partial x_1^j},\frac{\partial U}{\partial x_2^j},\frac{\partial U}{\partial x_3^j}\right).$$

证明这个二体问题的方程可以写成

$$m_j\boldsymbol{X}_j''=-\mathrm{grad}_j(U).$$

9. 证明该系统的总能量 $K+U$ 是一个运动常数, 其中

$$K=\frac{1}{2}\left(m_1|\boldsymbol{V}_1|^2+m_2|\boldsymbol{V}_2|^2\right).$$

10. 这个系统的角动量定义为

$$l=m_1(\boldsymbol{X}_1\times\boldsymbol{V}_1)+m_2(\boldsymbol{X}_2\times\boldsymbol{V}_2),$$

证明 l 也是一个首次积分.

第 14 章　洛伦茨系统

在我们至今研究的所有微分方程中，没有出现任何“混沌”. 理由是简单的：开始几章的线性系统总是展示简单可预测的行为. (是的，在第 6 章的振子系统中，我们看到解在环面上稠密缠绕，但这不是混沌.) 而且，对于后面几章中的非线性平面系统，庞加莱-本迪克逊定理完全排除了混沌行为的可能性. 所以，为了找到混沌行为，我们需要在非线性的高维系统中去寻找.

在本章中，我们研究气象学中的洛伦茨 (Lorenz) 系统，毫无疑问，这个系统在所有混沌微分方程中是最著名的. 这个系统首先由 E. N. 洛伦茨在 1963 年作为大气对流的一个极其简化模型而引入，它具有后来人们所熟知的“奇怪吸引子”. 在洛伦茨模型被人们知道之前，微分方程中的稳定吸引子已知类型仅有平衡点和闭轨. 在科学技术的所有领域，洛伦茨系统都真正为人们开辟了新的视野：后来发现洛伦茨系统中的许多现象在我们前面研究的所有领域 (生物学、电路理论、力学以及其他领域) 中都存在.

混沌系统的研究在随后的 40 年取得了极大的进展. 然而，需要提醒的是，分析像洛伦茨系统这样特定系统的混沌行为通常是极其困难的. 多数可以理解的混沌行为都来自于特定微分方程的几何模型，而不是实际方程本身. 事实上，这就是我们这里采取的方法. 我们将介绍洛伦茨系统的一个几何模型；利用离散动力学中的工具，可以完全地分析这个模型. 虽然 30 多年前人们就知道了这个模型，有趣的是，直到 1999 年，才证明这个模型与洛伦茨系统是等价的.

14.1　洛伦茨系统介绍

在 1963 年，E. N. 洛伦茨 [29] 试图建立一个微分方程系统来解释某些不可预测的天气行为. 多数实际的天气模型都涉及偏微分方程，洛伦茨找到了一个非常简单而又容易分析的系统.

洛伦茨模型可以如下 (多少有些不太精确地) 考虑. 想象一颗行星的“大气层”仅由一种单一的流体质点构成. 像地球上一样，质点在下面被加热 (然后上升) 而且在上面被冷却 (于是再落下). 气象专家能够预测这颗行星上的“天气”吗？可悲的是，回答是“不能”，这就对地球上准确预报天气产生了大量的问题，而地球大气层实际上有好多种质点.

准确点说，洛伦茨考虑一个二维流元，从它的下面加热而在上面冷却. 流体的

运动可以用一个包含无穷多个变量的微分方程系统来描述. 洛伦茨作了一个巨大的简化假设: 这些变量中除了 3 个外都保持常值. 粗略地说, 剩下的 3 个独立变量分别量度对流 "翻转" 速度 (x), 以及水平方向和竖直方向的温度变化 (分别为 y 和 z). 发生的运动就导致一个三维微分方程系统, 其中涉及 3 个参数: Prandtl 常数 σ, Rayleigh 常数 r 以及另外一个参数 b, 它与系统的实际大小有关. 当作了所有这些简化后, 最后的微分方程系统包含两个非线性项, 它的具体方程为

$$
\begin{aligned}
x' &= \sigma(y-x), \\
y' &= rx - y - xz, \\
z' &= xy - bz.
\end{aligned}
$$

在这个系统中, 假设 3 个参数都是正的, 而且 $\sigma > b+1$. 我们将这个系统记为 $\boldsymbol{X}' = \mathcal{L}(\boldsymbol{X})$. 当参数 $\sigma = 10, b = 8/3, r = 28$ 时, 我们在图 14.1 中画出了两条解曲线, 对应的初值条件分别为 $P_1 = (0,2,0)$ 和 $P_2 = (0,-2,0)$. 上面的参数就是洛伦茨发现中的原始参数. 注意, 两个解在出发时非常不同, 然而最终它们的命运多少是相同的: 它们似乎都围绕着一对点在缠绕, 而且时常变换所缠绕的对象. 这是关于洛伦茨系统的第一个重要事实: 所有非平衡解最终都趋向同一个复杂的集合, 也就是所谓的**洛伦茨吸引子**.

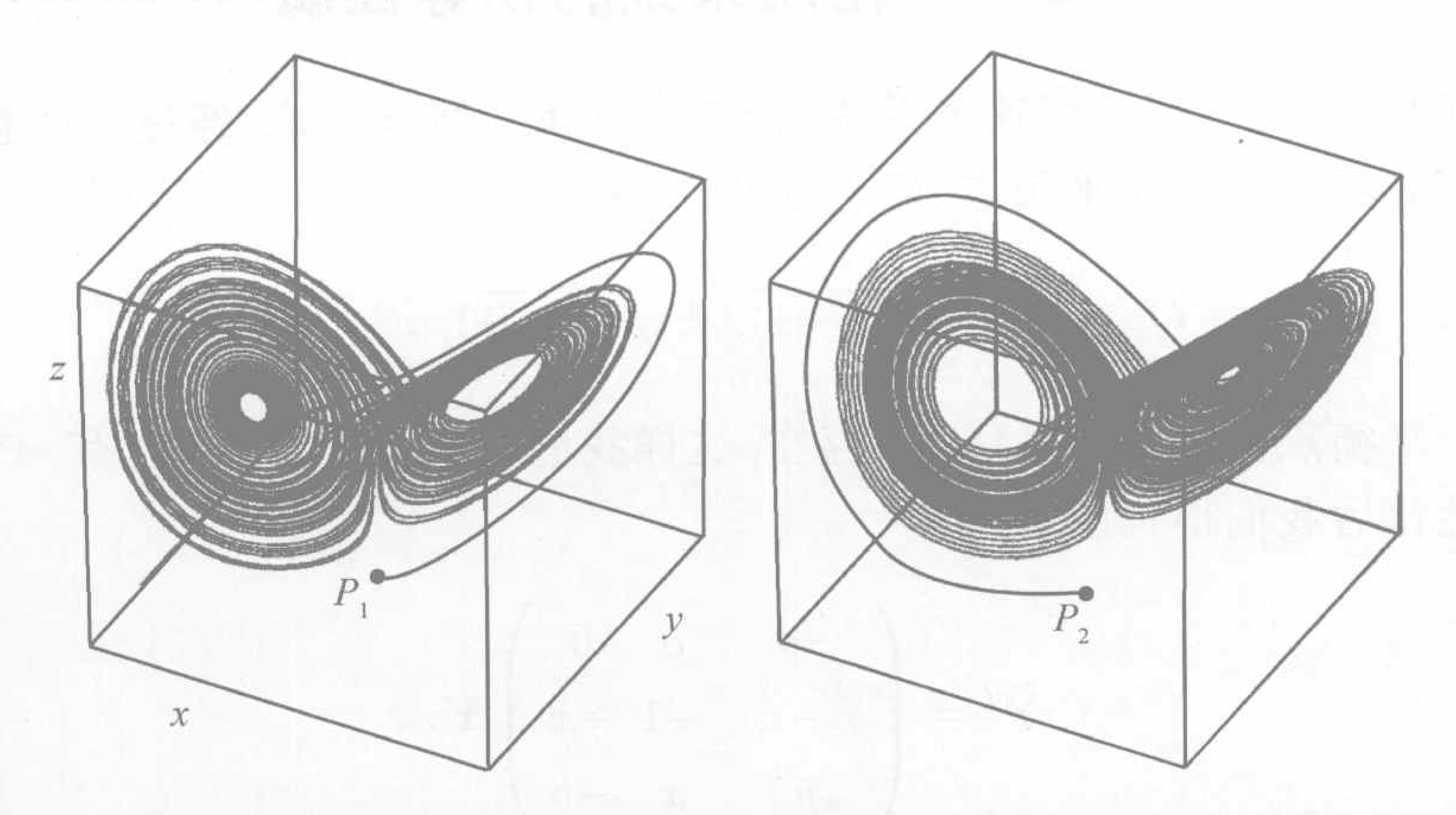


图 14.1 洛伦茨吸引子. 两个解分别对应于初值条件 $P_1 = (0,2,0)$ 和 $P_2 = (0,-2,0)$

这里隐藏着另一个重要事实. 在图 14.1 中, 我们从离开相对较远的两个点出发. 如果我们从两个很接近的初值条件出发, 将不会再观察到图 14.1 中明显的 "短暂行为". 这时, 对两个解, 我们将得到多少相同的图像. 但是, 这么说其实是在误导读者. 当画这两个解的实际坐标时, 我们发现, 在围绕洛伦茨吸引子的旅程中, 这两个解事实上相隔很远. 这点如图 14.2 所示, 其中我们画出了两个相邻出发的解的 x 坐标, 其中一个从 $(0,2,0)$ 出发, 另一个从 $(0,2.01,0)$ 出发. 在一定的时间

段, 这两个图像几乎重合, 然而当一个解绕着吸引子的一瓣前进, 而另一个解绕着另一瓣前进时, 这两个图像就相当不同了. 不管两个解出发时多么接近, 当它们逼近吸引子时, 它们总是按这种方式分开. 这就是**对初值条件的敏感依赖性**, 它是混沌系统的主要特征之一.

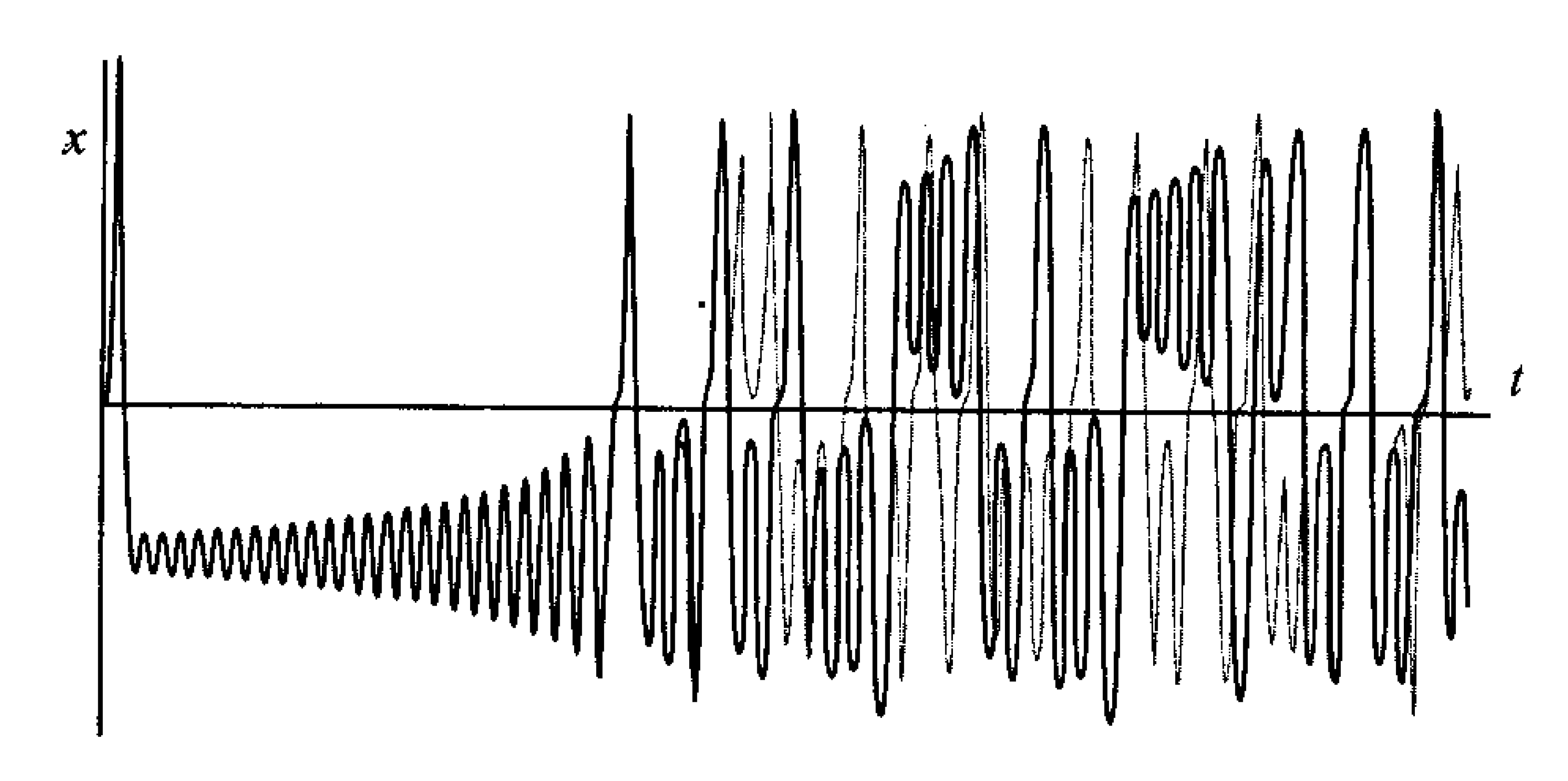


图 14.2 两个相邻初值条件 $P_1 = (0, 2, 0)$ 和 $P_2 = (0, 2.01, 0)$ 的 $x(t)$ 图像

我们将在本章中详细描述吸引子和混沌的概念. 但首先, 我们需要研究洛伦茨系统的一些常见特征.

14.2 洛伦茨系统的初等性质

像通常一样, 为了分析这个系统, 我们从寻找平衡点开始. 经过一些容易的代数运算可知系统有 3 个平衡点, 原点以及

$$Q_\pm = \left(\pm\sqrt{b(r-1)},\ \pm\sqrt{b(r-1)},\ r-1\right).$$

后面 2 个平衡点只有当 $r > 1$ 时才存在, 这样我们就看到在 $r = 1$ 时有一个分岔.

线性化后我们得到系统

$$\boldsymbol{Y}' = \begin{pmatrix} -\sigma & \sigma & 0 \\ r-z & -1 & -x \\ y & x & -b \end{pmatrix} \boldsymbol{Y}.$$

上面矩阵在原点处的特征值是 $-b$ 以及

$$\lambda_\pm = \frac{1}{2}\left(-(\sigma+1) \pm \sqrt{(\sigma+1)^2 - 4\sigma(1-r)}\right).$$

注意, 当 $0 \leqslant r < 1$ 时, $\lambda_\pm$ 都是负的. 从而, 原点此时是一个汇点.

洛伦茨系统 $\mathcal{L}(\boldsymbol{X})$ 具有一种对称性. 如果令 $S(x,y,z) = (-x,-y,z)$, 则我们有 $S(\mathcal{L}(\boldsymbol{X})) = \mathcal{L}(S(\boldsymbol{X}))$. 即, 关于 z 轴的反射保持这个向量场. 特别地, 如果

$(x(t), y(t), z(t))$ 是洛伦茨方程的一个解, 则 $(-x(t), -y(t), z(t))$ 也是该方程的一个解.

当 $x = y = 0$ 时, 有 $x' = y' = 0$, 从而 z 轴是不变的. 在这个轴上, 我们简单地有 $z' = -bz$, 因而这个轴上的所有的解都趋于原点. 事实上, 当 $r < 1$ 时, 过 $\mathbb{R}^3$ 中的任一点的解都趋于原点. 也就是:

命题 假设 $r < 1$. 洛伦茨系统的所有解都趋于位于原点的平衡点.

证明 我们来构造整个 $\mathbb{R}^3$ 上的一个严格李雅普诺夫函数. 令

$$L(x, y, z) = x^2 + \sigma y^2 + \sigma z^2.$$

于是我们有

$$\dot{L} = -2\sigma\big(x^2 + y^2 - (1 + r)xy\big) - 2\sigma b z^2.$$

于是, 如果在 $(x, y) \neq 0$ 时, 有

$$g(x, y) = x^2 + y^2 - (1 + r)xy > 0,$$

则在离开原点时就有 $\dot{L} < 0$. 沿 y 轴, 这显然正确. 沿 xy 平面上的其他直线 $y = mx$, 我们有

$$g(x, mx) = x^2(m^2 - (1 + r)m + 1).$$

但是当 $r < 1$ 时, 容易验证, 二项式 $m^2 - (1 + r)m + 1$ 对于所有的 m 都是正的. 因而当 $(x, y) \neq (0, 0)$ 时, 有 $g(x, y) > 0$. ■

当 r 增加并通过值 1 后, 要发生两件事. 首先, 原点处的特征值 λ_+ 变正, 于是原点现在是一个鞍点, 它具有二维的稳定曲面和一维的不稳定曲线. 其次, 在 $r = 1$ 时, 两个平衡点 $Q_\pm$ 在原点产生, 并且在 r 增加时, 它们从原点离开.

命题 如果

$$1 < r < r^* = \sigma\left(\frac{\sigma + b + 3}{\sigma - b - 1}\right),$$

则两个平衡点 $Q_\pm$ 都是汇点.

证明 从线性化我们可以算出 $Q_\pm$ 处的特征值满足三次多项式

$$f_r(\lambda) = \lambda^3 + (1 + b + \sigma)\lambda^2 + b(\sigma + r)\lambda + 2b\sigma(r - 1) = 0.$$

当 $r = 1$ 时, 多项式 f_1 有 3 个不同的根, 分别是 $0, -b$ 和 $-\sigma - 1$. 这些根不同是因为 $\sigma > b + 1$, 于是有

$$-\sigma - 1 < -\sigma + 1 < -b < 0.$$

从而当 r 大于 1 但靠近 1 时, f_r 有接近这 3 个值的 3 个实根. 注意, 当 $\lambda \geqslant 0$ 且 $r > 1$ 时, $f_r(\lambda) > 0$. 从 f_r 的图像可以看出, 至少当 r 接近 1 时, f_r 的 3 个根都是实的, 而且都为负.

我们现在让 r 增加, 来看 f_r 具有零实部特征值时 r 的最小取值. 注意, 由于 f_r 是一个实多项式而且当 $r>1$ 时它没有零根, 因而这个特征值的形式必定是 $\pm\mathrm{i}\omega$, $\omega\neq 0$. 通过在方程 $f_r(\mathrm{i}\omega)=0$ 中同时令虚部和实部为零就可得出命题的结论 (注意, 我们已经假设 $\sigma>b+1$). ■

人们已经知道, 在 r^* 处会出现一个霍普夫分岔, 但是证明这件事超出了本书的范围.

当 $r>1$ 时, 所有解都趋于原点这件事不再正确. 但是, 我们可以说, 从远离原点的地方出发的解的确会运动到较靠近的地方. 为了说得准确, 我们令

$$V(x,y,z)=rx^2+\sigma y^2+\sigma(z-2r)^2.$$

注意, 方程 $V(x,y,z)=v>0$ 定义了 $\mathbb{R}^3$ 中一个中心在 $(0,0,2r)$ 处的椭球面. 我们来证明:

命题 存在 v^* 使得所有从椭球 $V\leqslant v^*$ 外部出发的解, 最终都要进入这个椭球, 并且在将来的所有时间都始终陷在这个椭球中.

证明 我们来作计算

$$\begin{aligned}\dot V &= -2\sigma\big(rx^2+y^2+b(z^2-2rz)\big)\\ &= -2\sigma\big(rx^2+y^2+b(z-r)^2-br^2\big).\end{aligned}$$

方程

$$rx^2+y^2+b(z-r)^2=\mu$$

在 $\mu>0$ 时也定义了一个椭球面. 当 $\mu>br^2$ 时, 我们有 $\dot V<0$. 于是, 我们可取 v^* 充分大使得椭球面 $V=v^*$ 的内部严格包含椭球面

$$rx^2+y^2+b(z-r)^2=br^2.$$

则当 $v\geqslant v^*$ 时, 就有 $\dot V<0$. ■

作为一个推论, 我们得出, 所有从远离原点出发的解都被吸引到位于椭球面 $V=v^*$ 内部的一个集合. 记 Λ 为所有这样的点 P 构成的集合, 从 P 出发的解在所有时间 (向前和向后) 都始终逗留在椭球面 $V=v^*$ 的内部. 于是洛伦茨系统的任一解的 ω 极限集都必定位于 Λ 中. 理论上, Λ 可以是一个很大的集合, 有可能包含 $\mathbb{R}^3$ 中的一个开区域. 然而, 对于洛伦茨系统, 情况不是这样的.

为了看出这点, 从微积分中我们知道, $\mathbb{R}^3$ 中的一个向量场 $\boldsymbol{F}(\boldsymbol{X})$ 的**散度**定义为

$$\operatorname{div}\boldsymbol{F}=\sum_{i=1}^{3}\frac{\partial \boldsymbol{F}}{\partial x_i}(\boldsymbol{X}).$$

$\boldsymbol{F}$ 的散度量度体积在 $\boldsymbol{F}$ 的流 ϕ_t 作用下的变化快慢. 假设 D 是 $\mathbb{R}^3$ 中一个具有光滑边界的区域, 令 $D(t) = \phi_t(D)$, 即 D 在流的时间 t 映射下的像. 记 $V(t)$ 为 $D(t)$ 的体积. 刘维尔定理断言

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \int_{D(t)} \operatorname{div} F \, \mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

对于洛伦茨系统, 我们可以立刻算出它的散度为常数 $-(\sigma+1+b)$, 因而体积按常值速率减少

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -(\sigma+1+b)V.$$

求解这个简单的微分方程就可得出

$$V(t) = \mathrm{e}^{-(\sigma+1+b)t}V(0),$$

因而任何体积都按指数速度缩小到 0. 特别地, 我们有:

命题 Λ 的体积为零. ■

一个自然的问题就是, 对于"吸引子" Λ 的结构, 我们还能说些什么呢? 在二维时, 这样的一个集合由一些极限环、平衡点以及连接它们的解构成. 在高维时, 这些吸引子可以非常"奇怪", 我们将在下节看出这点.

14.3 洛伦茨吸引子

洛伦茨系统在参数 r 增加时的行为是非常现代的研究领域, 严格了解这个系统在参数变化时出现的令人着迷的所有动力现象可能还需要几十年 (不说几百年的话). 关于这个领域, Sparrow[44] 已经写了一整本书.

在本节我们将讨论一个特别的参数集, 此时洛伦茨系统有一个吸引子. 粗略地说, 流的一个吸引子就是一个"吸引"所有附近解的不变集. 准确地说就是:

定义 设 $\boldsymbol{X}' = F(\boldsymbol{X})$ 是 $\mathbb{R}^n$ 中的一个微分方程系统, 它产生的流记为 ϕ_t. 一个集合 Λ 称为一个**吸引子**, 如果

(1) Λ 是紧的和不变的;

(2) 存在包含 Λ 的开集 U 使得对每个 $\boldsymbol{X} \in U$ 和所有的 $t \geqslant 0$, 都有 $\phi_t(\boldsymbol{X}) \in U$, 而且 $\cap_{t\geqslant 0}\phi_t(U) = \Lambda$;

(3) (传递性) 任给两个点 $Y_1, Y_2 \in \Lambda$, 以及 Y_j 在 U 中的任意开邻域 U_j, 都存在从 U_1 中出发的一条解曲线经过 U_2. ■

定义中的传递性条件似乎有些奇怪. 基本上, 我们把它包含进来是为了保证我们寻找的只是一个吸引子, 而不是一组动力学不同的吸引子. 例如, 我们来看一个平面系统的例子, 传递性条件会排除掉这种情形:

$$x' = x - x^3,$$
$$y' = -y.$$

这个系统的相图如图 14.3 所示. 注意, 该系统的任何解都要进入图中标示为 U 的集合, 然后再趋于以下 3 个平衡点之一: 位于 $(\pm1,0)$ 的汇点以及位于 $(0,0)$ 的鞍点. 流 ϕ_t 在 U 上作用后的正向交集为 x 轴上的区间 $-1 \leqslant x \leqslant 1$. 这个区间满足定义中的条件 1 和条件 2, 但与条件 3 冲突, 这是因为没有解曲线同时经过这个区间的左端点和右端点. 我们倾向于认为这个集合不是一个吸引子, 因为多数解都趋于 2 个汇点中的一个. 在这种情形, 我们实际上有 2 个不同的吸引子.

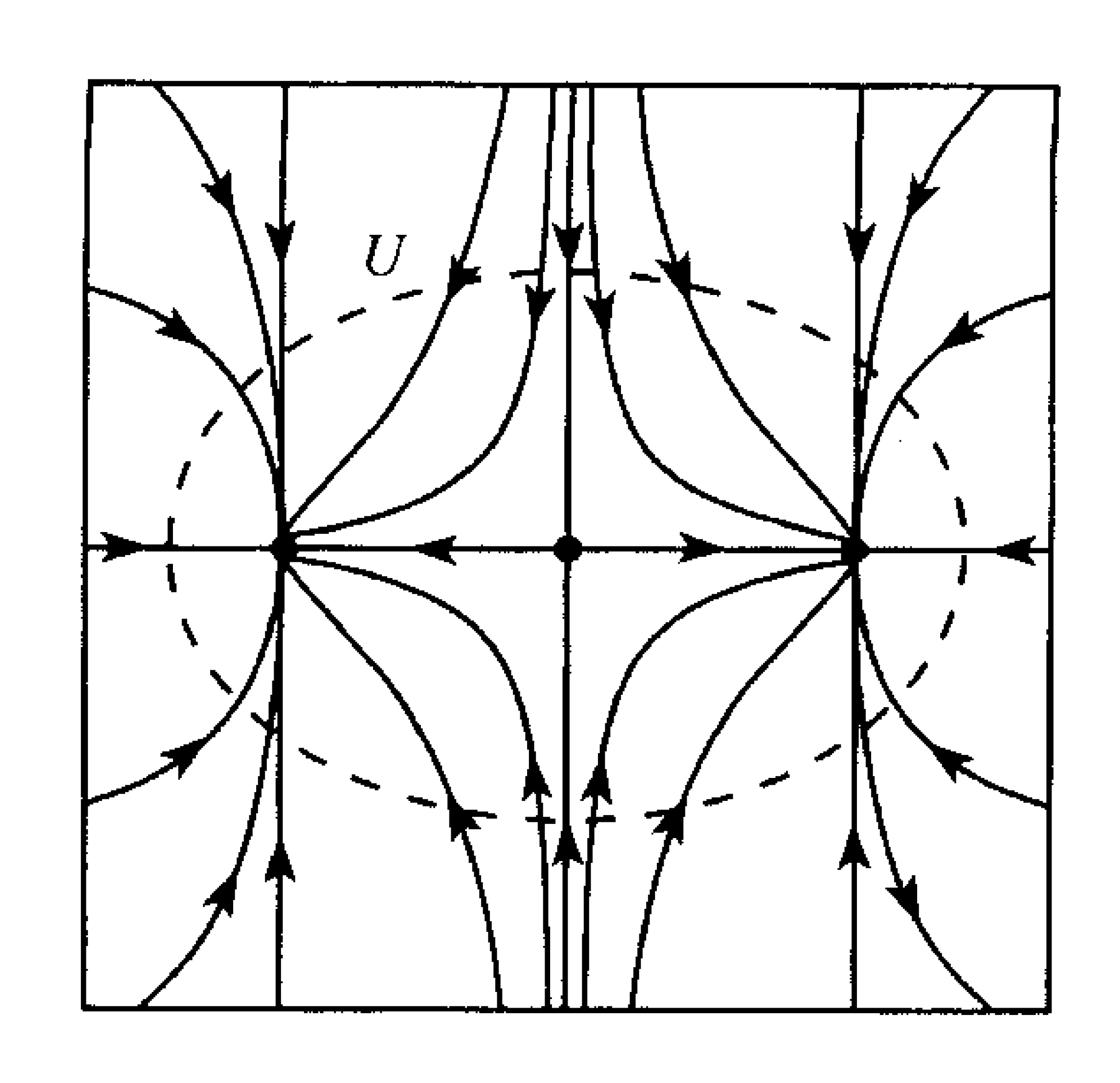


图 14.3 虽然所有的解都进入 U, 但是 x 轴上位于两个汇点之间的区间并不是该系统的一个吸引子

需要说明的是, 数学上没有一个被普遍接受的吸引子定义. 有些人倾向于认为满足条件 1 和条件 2 的集合 Λ 就是一个吸引子, 而当 Λ 也满足条件 3 时, 它就被称为一个传递的吸引子. 对平面系统, 条件 3 常常很容易验证; 但是在高维情形, 下面我们将看到, 这点可能会非常困难.

在本章剩下的部分, 我们将注意力集中到洛伦茨系统一个非常特殊的情形: 此时的参数分别为 $\sigma = 10, b = 8/3, r = 28$. 历史上, 洛伦茨首先使用这些参数在这个系统中观察到了混沌现象. 于是, 我们所考虑的特定洛伦茨系统就是

$$\boldsymbol{X}' = \mathcal{L}(\boldsymbol{X}) = \begin{pmatrix} 10(y-x) \\ 28x - y - xz \\ xy - (8/3)z \end{pmatrix}.$$

同上节一样, 我们有 3 个平衡点: 原点以及 $Q_\pm = (\pm6\sqrt{2}, \pm6\sqrt{2}, 27)$. 在原点我们找到特征值 $\lambda_1 = -8/3$ 和

$$\lambda_\pm = -\frac{11}{2} \pm \frac{\sqrt{1201}}{2}.$$

为了后面的应用, 需要注意这些特征值满足

$$\lambda_- < -\lambda_+ < \lambda_1 < 0 < \lambda_+.$$

于是原点处的线性化系统就是

$$\boldsymbol{Y}' = \begin{pmatrix} \lambda_- & 0 & 0 \\ 0 & \lambda_+ & 0 \\ 0 & 0 & \lambda_1 \end{pmatrix} \boldsymbol{Y}.$$

这个线性化系统的相图如图 14.4 所示. 注意, 该系统的稳定平面上的所有解 (x 轴上的解除外) 都是切于 z 轴地趋于原点.

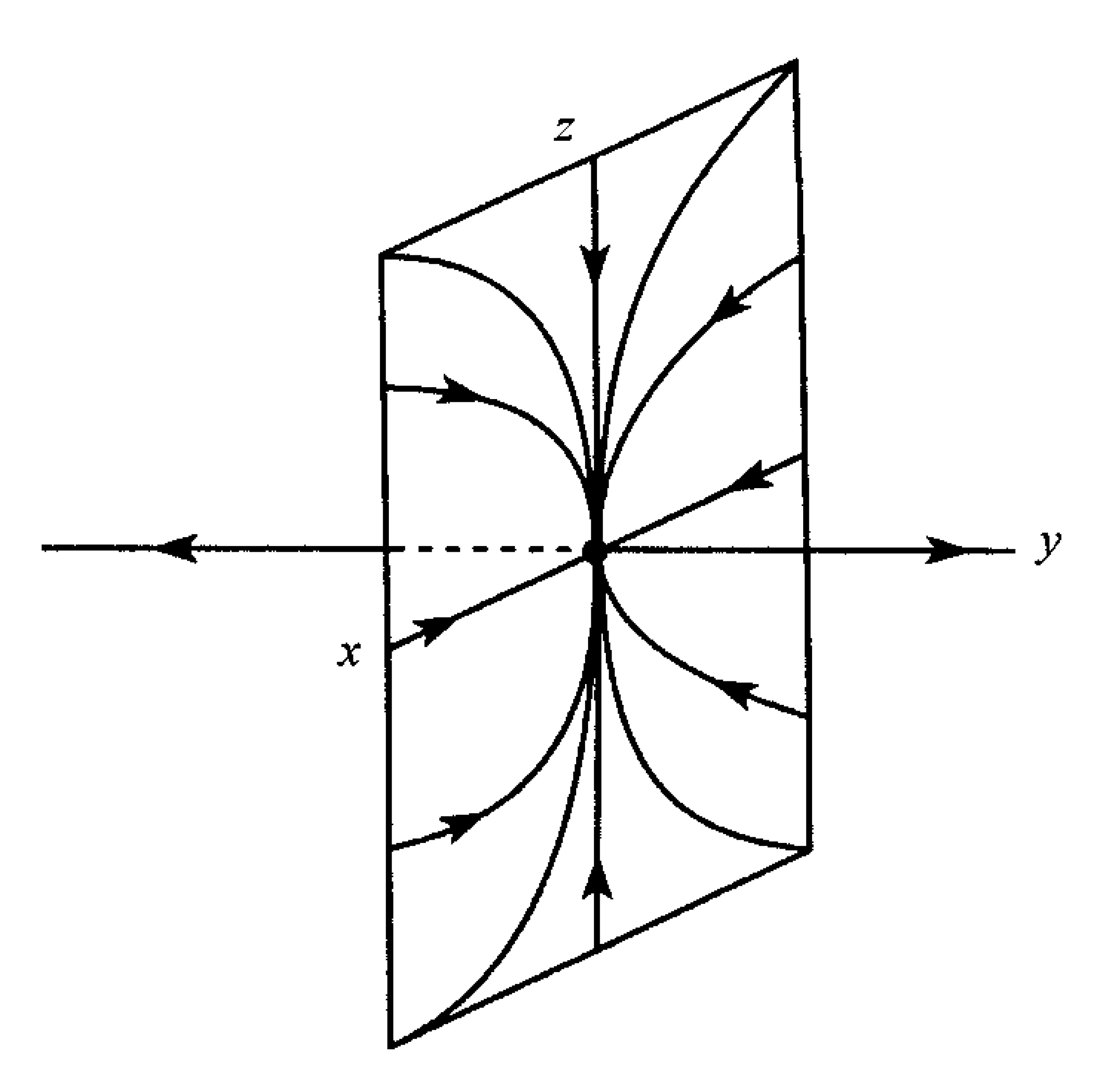


图 14.4 洛伦茨系统在原点处的线性化

在 $Q_\pm$ 处, 计算表明仅有一个负实特征值和一对具有正实部的复共轭特征值. 注意, 洛伦茨系统中的对称性迫使关于 Q_+ 和 Q_- 的旋转方向相反.

在图 14.5 中, 我们展示了通过数值计算得到的原点不稳定曲线左右两支的一部分. 注意, 不稳定曲线的右边一支先是接近 Q_-, 然后再盘旋地离开. 左边一支的行为 (与右边一支的行为) 关于过 z 轴的反射对称. 在图 14.6 中, 我们展示了不稳定曲线引人注目的很大部分. 注意, 它们似乎在围绕这 2 个平衡点旋转, 有时绕着 Q_- 盘旋, 有时绕着 Q_+ 盘旋. 特别地, 这些曲线与平面 $z = 27$ 上包含 $Q_\pm$ 的部分连续不断地相交, 在平面 $z = 27$ 上的这一部分, 向量场指向下方. 这使我们想到在平面 $z = 27$ 的这一部分上构造一个庞加莱映射. 正如我们在前面看到的那样, 计算庞加莱映射通常是不可能的, 现在的情形也不例外. 因而, 我们将仅满足于建立一个简化的模型, 这个模型具有我们在洛伦茨系统中发现的许多行为. 我们在接

下来的一节中将看到, 这一模型提供了一个可计算的手段, 使得我们可以估算该系统的混沌行为.

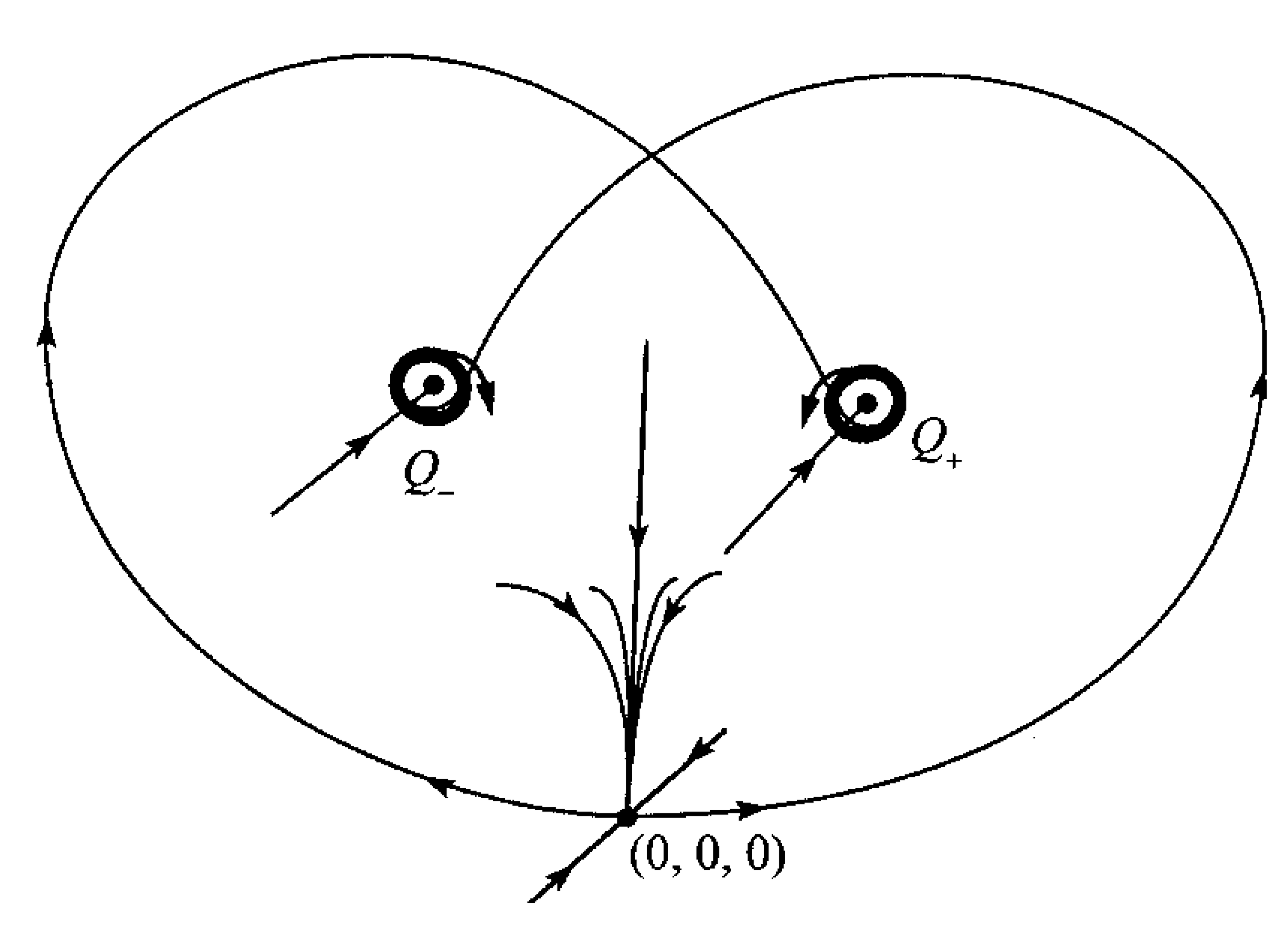


图 14.5 原点的不稳定曲线

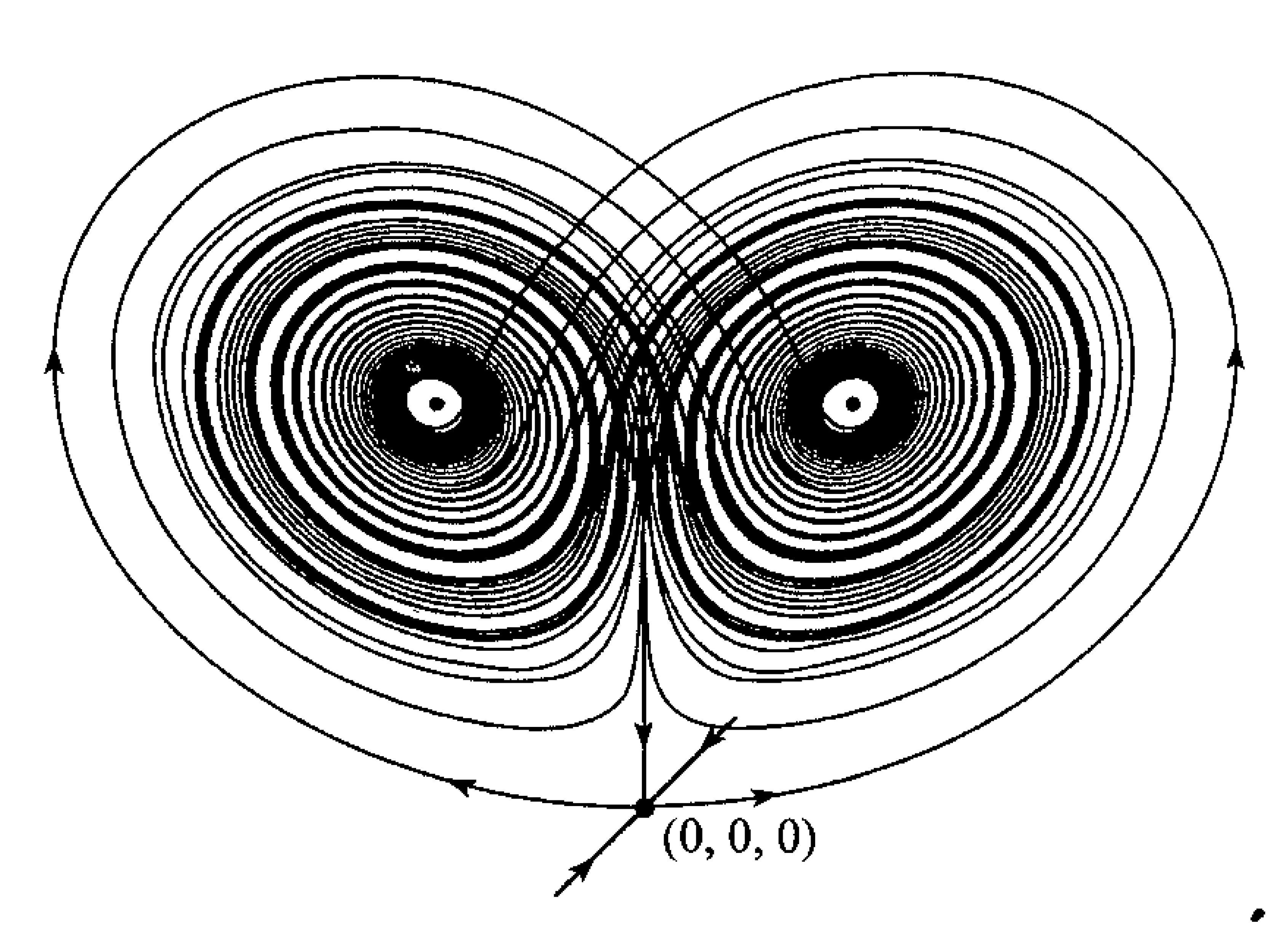


图 14.6 原点不稳定曲线的更多部分

14.4 洛伦茨吸引子的一个模型

在本节, 我们描述洛伦茨吸引子的一个几何模型, 它最初由 Guckenheimer 和 Williams[20] 提出. Tucker[46] 证明了, 对于某些参数, 这个模型的确真正对应于原来的洛伦茨系统. 我们并不明确地指定该模型的向量场, 而是给它的流一个定性描述, 这一点非常类似于我们曾在第 11 章中采取的方式. 我们所使用的一些特定数字并不重要, 重要的是它们之间的相对大小.

与洛伦茨系统一样, 我们假设这个模型关于反射 $(x,y,z) \to (-x,-y,z)$ 是对称的. 我们首先在 $\mathbb{R}^3$ 的原点放置一个平衡点, 并且假设在由 $|x|, |y|, |z| \leqslant 5$ 所给出的立方体 $\mathcal{S}$ 中, 系统是线性的. 我们不使用实际洛伦茨系统中的特征值 λ_1 和 $\lambda_\pm$,

为了简化一点计算, 我们假设特征值为 −1, 2 和 −3, 从而在立方体 $\mathcal{S}$ 中, 模型的系统如下给出：

$$\begin{aligned} x' &= -3x, \\ y' &= 2y, \\ z' &= -z. \end{aligned}$$

注意, 这个系统的相图与图 14.4 中给出的相图一致, 而且特征值之间的相对大小与洛伦茨系统情形也是相同的.

我们需要知道解在 (0,0,0) 附近是如何经过的. 考虑平面 $z=1$ 中由 $|x| \leqslant 1, 0<y\leqslant \epsilon<1$ 所给出的矩形 $\mathcal{R}_1$. 当时间向前时, 所有从 $\mathcal{R}_1$ 中出发的解最终都要到达矩形 $\mathcal{R}_2$, 其中矩形 $\mathcal{R}_2$ 位于平面 $y=1$ 中, 由 $|x|\leqslant 1, 0<z\leqslant 1$ 所定义. 于是, 我们就得到一个函数 $h:\mathcal{R}_1\to\mathcal{R}_2$, 跟随解曲线从 $\mathcal{R}_1$ 到 $\mathcal{R}_2$ 就给出了 h 的定义. 我们可以验证 (这一验证留作习题) 这个函数具有形式

$$h\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} xy^{3/2} \\ y^{1/2} \end{pmatrix}.$$

由此可知, h 将 $\mathcal{R}_1$ 中的线段 $y=c$ 带到 $\mathcal{R}_2$ 的线段 $z=c^{1/2}$. 同时, 由于 $x_1=xz_1^3$, 可知 h 将线段 $x=c$ 映到形如 $x_1=cz_1^3$ 的曲线. 每条这种像曲线都与 xy 平面垂直相交, 如图 14.7 所示.

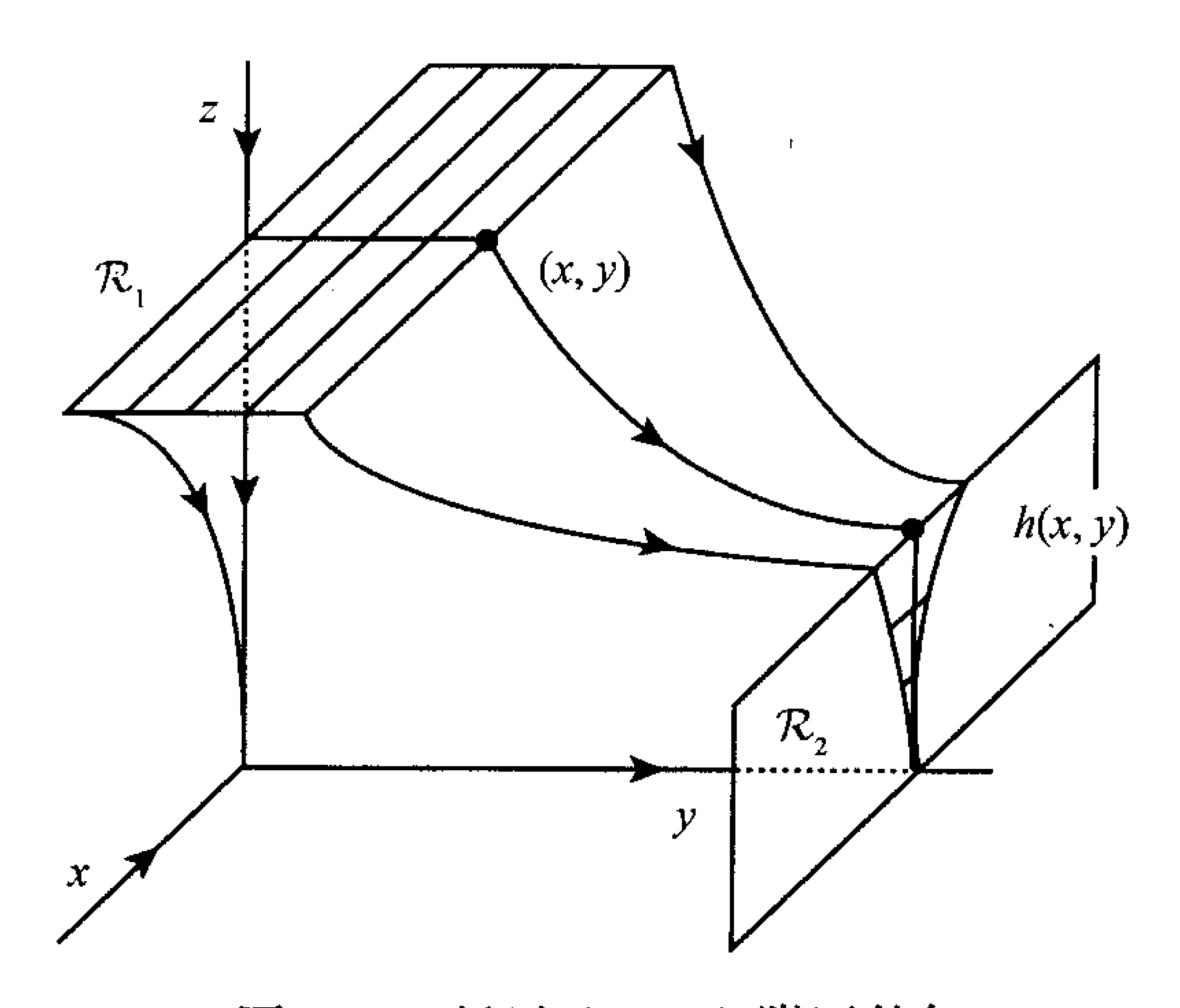


图 14.7 经过 (0,0,0) 附近的解

模仿洛伦茨系统, 我们在平面 $z=27$ 上放置另外 2 个平衡点：一个位于 $Q_-=(-10,-20,27)$, 而另一个位于 $Q_+=(10,20,27)$. 我们假设由 $y=\pm 20, z=27$ 给出的线段构成 $Q_\pm$ 的稳定直线的一部分, 而且这些平衡点处的另外 2 个特征值是具有正实部的复数.

记 Σ 为正方形 $|x|,|y| \leqslant 20, z=27$. 我们假设在这个正方形的内部向量场指向下方. 从而解按与洛伦茨系统相同的方式盘旋离开 $Q_{\pm}$. 我们还假设 $(0,0,0)$ 的稳定曲面与 Σ 首先相交于 xz 平面和 Σ 的交线上.

记 $\zeta^{\pm}$ 为原点不稳定曲线的 2 支. 我们假设这些曲线先绕着 Σ 走一段, 然后再进入这个正方形, 如图 14.8 所示. 我们记 $\zeta^{\pm}$ 与 Σ 的第一个交点为 $\rho^{\pm}=(\pm x^*, \mp y^*)$.

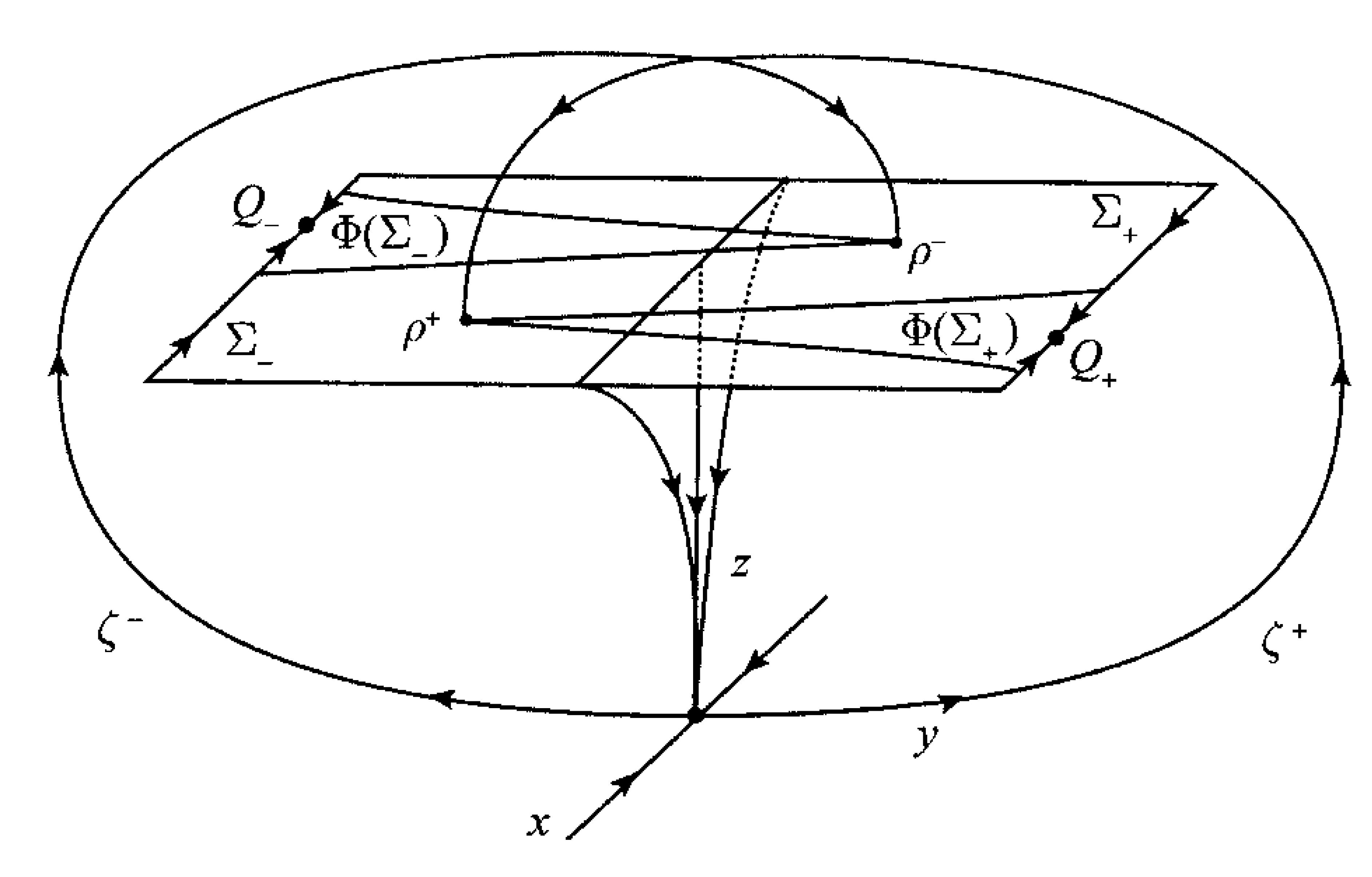


图 14.8 洛伦茨吸引子模型中的解 $\zeta^{\pm}$ 以及它们与 Σ 的交点

现在考虑 Σ 中的一条线段 $y=v$. 如果 $v=0$, 当时间向前时, 所有从这条线段上的点出发的解都趋于原点. 于是, 这些解将不再返回 Σ. 我们假设, 当时间向前时, 所有其他从 Σ 出发的解都要返回 Σ. 我们关于这个模型的主要假设就是这些解是如何返回的:

(1) **返回条件**: 令 $\Sigma_+ = \Sigma \cap \{y>0\}$, 而 $\Sigma_- = \Sigma \cap \{y<0\}$. 我们假设当时间向前时, 从 $\Sigma_{\pm}$ 中出发的所有解都返回 Σ. 于是, 我们就得到一个庞加莱映射 $\Phi: \Sigma_+ \cup \Sigma_- \to \Sigma$. 我们假设像 $\Phi(\Sigma_{\pm})$ 如图 14.8 所示. 根据对称性, 我们有 $\Phi(x,y) = -\Phi(-x,-y)$.

(2) **压缩方向**: 我们假设存在某个函数 g, 使得对每个 $v \neq 0$, Φ 将 Σ 中的线段 $y=v$ 映到线段 $y=g(v)$. 我们进一步假设 Φ 沿 x 方向压缩这些线段.

(3) **扩张方向**: 我们假设 Φ 沿 y 方向将 Σ_- 和 Σ_+ 拉长一个严格大于 $\sqrt{2}$ 的系数, 于是有 $g'(y) > \sqrt{2}$.

(4) **双曲性条件**: 除了扩张和压缩条件外, 我们还假设 $D\Phi$ 将切于 $\Sigma_{\pm}$ 的, 斜率为 ± 1 的向量映到斜率的大小大于 $\mu > 1$ 的向量.

分析上, 这些条件意味着映射 Φ 具有如下形式

$$\Phi(x,y) = (f(x,y),\ g(y)),$$

其中 $g'(y) > \sqrt{2}$, 且 $0 < \partial f/\partial x < c < 1$. 双曲性条件蕴涵

$$g'(y) > \mu \left| \frac{\partial f}{\partial x} \pm \frac{\partial f}{\partial y} \right|.$$

几何上, 这个条件意味着切平面上由 $|y| \geqslant |x|$ 给出的扇形被 $D\Phi$ 映到一个斜率更陡的扇形内部. 注意, 如果 $|\partial f/\partial y|$ 和 c 在整个 $\Sigma_\pm$ 上都充分小, 这个条件就能满足.

技术上, $\Phi(x,0)$ 没有定义, 但我们确有

$$\lim_{y\to 0^\pm} \Phi(x,y) = \rho^\pm,$$

注意, 这里的 $\rho^\pm$ 就是 $\zeta^\pm$ 与 Σ 的第一个交点. 我们称 $\rho^\pm$ 为 $\Phi(\Sigma_\pm)$ 的**尖端**. 事实上, 我们关于特征值的假设保证了当 $y \to 0$ 时, 有 $g'(y) \to \infty$(见本章的习题 3).

为了找到吸引子, 我们将注意力集中到由 $|y| \leqslant y^*$ 给出的矩形 $R \subset \Sigma$, 注意, 这里 $\mp y^*$ 就是尖端 $\rho^\pm$ 的 y 坐标. 令 $R_\pm = R \cap \Sigma_\pm$. 容易验证所有从 $\Sigma_\pm$ 的内部但 R 的外部出发的解最终都要与 R 相交, 因而只需考虑 Φ 在 R 上的行为. 图 14.9 是 Φ 在 R 上作用的一个平面示意图. 注意, $\Phi(R) \subset R$.

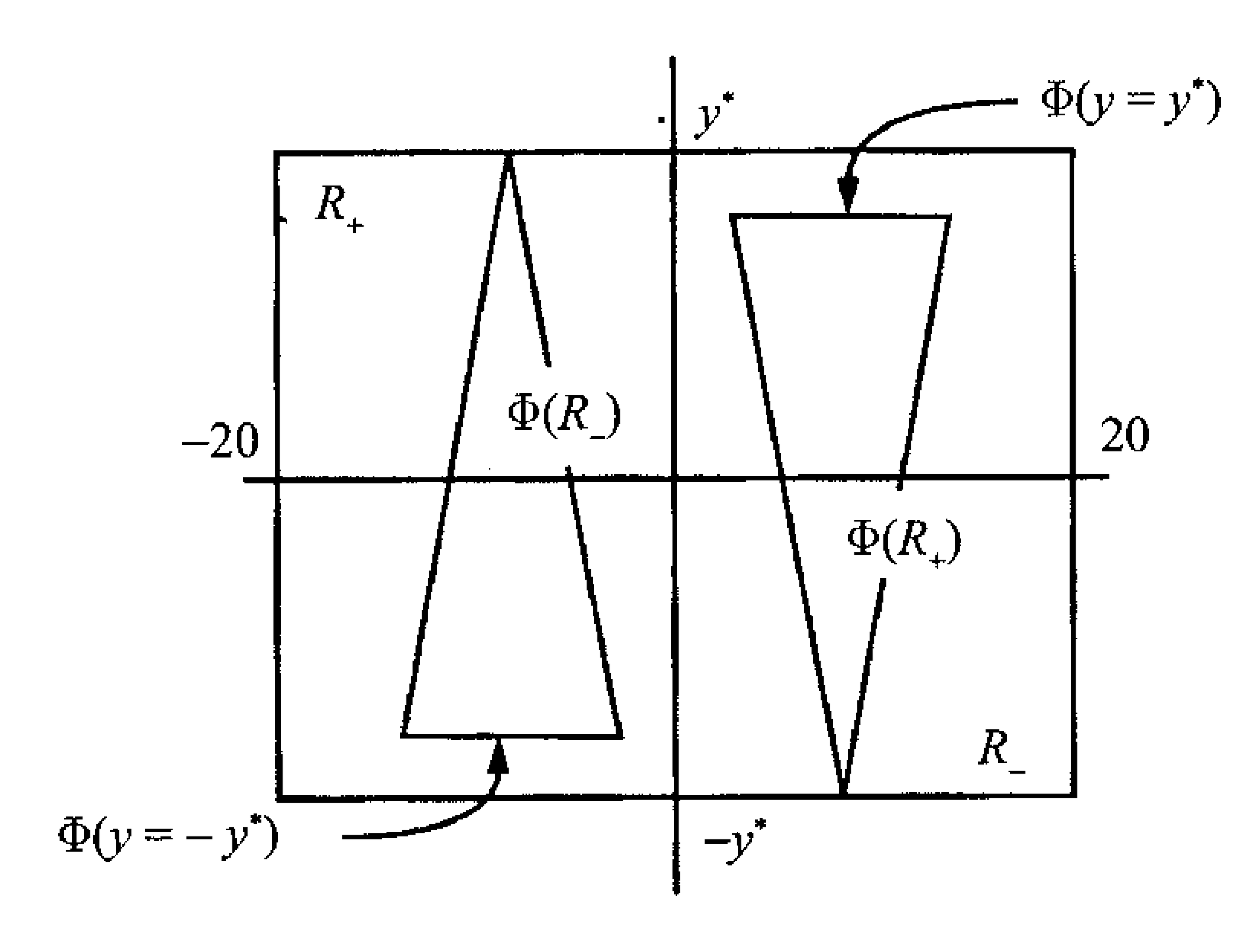


图 14.9 R 上的庞加莱映射 Φ

记 Φ^n 为 Φ 的 n 次迭代, 然后令

$$A = \bigcap_{n=0}^{\infty} \overline{\Phi^n(R)}.$$

这里 $\overline{U}$ 代表集合 U 的闭包. 集合 A 就将是这个流的吸引子与 R 的交集. 即, 令

$$\mathcal{A} = \left(\bigcup_{t\in\mathbb{R}} \phi_t(A) \right) \cup \{(0,0,0)\}.$$

我们加上原点是为了使得 $\mathcal{A}$ 成为一个闭集. 我们来证明:

定理　$\mathcal{A}$ 是模型洛伦茨系统的一个吸引子.

证明　只要证明了 A 是映射 Φ 的一个吸引子 (其中映射吸引子的定义与流的完全类似) 就立刻可以推出 $\mathcal{A}$ 是这个流的吸引子. 显然 A 是闭的. 技术上, A 本身在 Φ 之下并不是不变的, 这是因为沿 $y=0$, Φ 没有定义. 但是, 对于流 ϕ_t, 所有经过这种点的解都位于 $\mathcal{A}$ 中, 从而 $\mathcal{A}$ 是不变的. 如果我们令 $\mathcal{O}$ 是由 Σ 的内部给出的开集, 则对所有的 $(x,y)\in\mathcal{O}$, 都存在一个 n 使得 $\Phi^n(x,y)\in R$. 于是,

$$\bigcap_{n=0}^{\infty}\Phi^n(\mathcal{O})\subset A.$$

又, 根据定义, $A=\cap_{n\geqslant 0}\Phi^n(R)$, 从而也有 $A=\cap_{n\geqslant 0}\Phi^n(\mathcal{O})$. 因此, 吸引子定义中的条件 1 和条件 2 对 Φ 是成立的.

剩下需要证明的是传递性. 我们需要证明, 如果 P_1,P_2 是 A 中的两个点, 且 W_j 是 P_j 在 $\mathcal{O}$ 中的开邻域, 则存在 $n\geqslant 0$ 使得 $\Phi^n(W_1)\cap W_2\neq\emptyset$.

给定集合 $U\subset R$, 记 $\Pi_y(U)$ 为 U 在 y 轴上的投影. 同时, 记 $l_y(U)$ 为 $\Pi_y(U)$ 的长度, 我们称之为 U 的 y 长度. 在下面, U 将是有限个连通集合的并, 从而 $l_y(U)$ 是定义好的.

我们需要一个引理. ■

引理　任给开集 $W\subset R$, 存在 $n>0$ 使得 $\Pi_y(\Phi^n(W))$ 就是区间 $[-y^*,y^*]$. 换句话说就是, $\Phi^n(W)$ 与 R 中的每条线段 $y=c$ 都相交.

证明　首先假设 W 包含一个连通子集 W', 从 $y=0$ 延伸到其中一个尖端. 于是, $l_y(W')=y^*$. 因而 $\Phi(W')$ 是连通的, 而且有 $l_y(\Phi(W'))\geqslant\sqrt{2}y^*$. 进一步, $\Phi(W')$ 也延伸到其中一个尖端, 但是由于它的 y 长度超过了 y^*, 因而它要与 $y=0$ 相交.

现在用 Φ 再作用一次. 注意, $\Phi^2(W')$ 包含 2 片, 其中一片延伸到 ρ^+, 另一片延伸到 ρ^-. 而且 $l_y(\Phi^2(W'))>2y^*$. 由此可得 $\Pi_y(\Phi^2(W'))=[-y^*,y^*]$, 这样, 在这种特殊情形, 我们就完成了证明.

对于一般情形, 首先假设 W 是连通的而且不穿过 $y=0$. 于是与上面一样我们有 $l_y(\Phi(W))>\sqrt{2}l_y(W)$, 从而 $\Phi(W)$ 的 y 长度以超过 $\sqrt{2}$ 的倍数增加.

如果 W 真的穿过 $y=0$, 则我们可以找到一对连通集合 $W^{\pm}$ 使得 $W^{\pm}\subset R^{\pm}\cap W$ 且 $l_y(W^+\cup W^-)=l_y(W)$. 像 $\Phi(W^{\pm})$ 延伸到尖端 $\rho^{\pm}$. 如果这 2 个集合中的一个也与 $y=0$ 相交, 则根据上面, 我们也就完成了证明. 如果 $\Phi(W^+)$ 和 $\Phi(W^-)$ 都不与 $y=0$ 相交, 则我们可以再用 Φ 作用. $\Phi(W^+)$ 和 $\Phi(W^-)$ 都是连通集合, 而且我们有 $l_y(\Phi^2(W^{\pm}))>2l_y(W^{\pm})$. 因而, 对于 W^+ 或 W^- 中的一个, 我们有 $l_y(\Phi^2(W^{\pm}))>l_y(W)$, 从而 W 的 y 长度在迭代下也会增加.

于是, 如果我们继续迭代 $\Phi,\Phi^2,\cdots$, 并且在每一步都像上面那样选择 $\Phi^j(W)$ 的 y 长度最大的子集, 则我们就看到这些像的 y 长度将无限增加. 这就完成了引理

的证明. ■

现在我们来完成定理的证明. 我们必须在 W_1 中找到一点使得它在 Φ 的某次迭代下的像位于 W_2 中. 为此, 注意到,

$$\left|\Phi^k(x_1,y)-\Phi^k(x_2,y)\right|\leqslant c^k|x_1-x_2|,$$

这是因为对每个 j, $\Phi^j(x_1,y)$ 和 $\Phi^j(x_2,y)$ 都位于平行于 x 轴的同一直线上, 而且 Φ 以因子 $c<1$ 压缩 x 方向上的距离.

我们不妨假设 W_2 是一个半径为 ϵ 的圆盘. 注意, R 在 x 方向上的宽度为 40, 我们可取 m 使得 $40c^m<\epsilon$. 考虑 $\Phi^{-m}(P_2)$. 由于 $P_2\in\cap_{n\geqslant 0}\Phi^n(R)$, 因此 $\Phi^{-m}(P_2)$ 有定义. 令 $\Phi^{-m}(P_2)=(\xi,\eta)$.

从上一引理可知, 存在 n 使得

$$\Pi_y(\Phi^n(W_1))=[-y^*,y^*].$$

于是我们可以选取一点 $(\xi_1,\eta)\in\Phi^n(W_1)$. 设 $(\xi_1,\eta)=\Phi^n(\tilde{x},\tilde{y})$, 其中 $(\tilde{x},\tilde{y})\in W_1$. 由于 $\Phi^n(\tilde{x},\tilde{y})$ 与 $\Phi^{-m}(P_2)$ 具有相同的 y 坐标, 因而有

$$\begin{aligned}|\Phi^{m+n}(\tilde{x},\tilde{y})-P_2| &= |\Phi^m(\xi_1,\eta)-P_2|\\ &= |\Phi^m(\xi_1,\eta)-\Phi^m(\xi,\eta)|\\ &\leqslant 40c^m<\epsilon.\end{aligned}$$

这样我们就找到了一点 $(\tilde{x},\tilde{y})\in W_1$ 满足从该点出发的解经过 W_2. 这就完成了证明. ■

注意, 在上面的证明中, 从 P_1 附近出发且接近 P_2 的解曲线不一定位于这个吸引子中. 但是, 有可能真的在 $\mathcal{A}$ 中找到这样的一个解 (见习题 4).

14.5 混沌吸引子

在上节中, 我们将关于洛伦茨系统解的行为的研究约化成分析 Φ 的动力学. 在这个过程中, 我们从一个三维微分方程系统降维到了一个二维映射. 但是我们可以做得更好. 根据假设, Σ 中 2 个 y 坐标相同的点被映到了 2 个 y 坐标为 $g(y)$ 的新点, 于是 2 个新点的 y 坐标也相同. 而且, 这些点之间的距离被压缩了. 因而, 在 Φ 的迭代下, 我们不用关心线段 $y=$ 常数上的全部点; 我们只需跟踪 g 的迭代下的 y 坐标就可. 于是, 正如我们下面将要看到的那样, 庞加莱映射 Φ 由定义在区间 $[-y^*,y^*]$ 上的一维函数 g 的动力学完全决定. 事实上, 这个一维函数的迭代完全确定了吸引子中所有解的行为. 在本节, 我们开始研究**一维离散动力系统**的动力学. 在第 15 章, 我们将更深入地讨论这个课题.

记 I 为区间 $[-y^*,y^*]$. 注意, 除了 $y=0$ 外, g 在 I 上有定义, 而且满足 $g(-y)=-g(y)$. 从上节的结果可知, $g'(y)>\sqrt{2}$, 而且有 $0<g(y^*)<y^*$ 及 $-y^*<g(-y^*)<0$. 而且还有,

$$\lim_{y\to 0^\pm} g(y)=\mp y^*.$$

从而 g 的图像类似于图 14.10. 注意, 区间 $[g(-y^*),g(y^*)]$ 中的点都有 2 个原像, 而区间 $(-y^*,g(-y^*))$ 和 $(g(y^*),y^*)$ 中的点都只有一个原像. I 的端点, 即 $\pm y^*$ 在 I 中没有原像, 这是因为 $g(0)$ 没有定义.

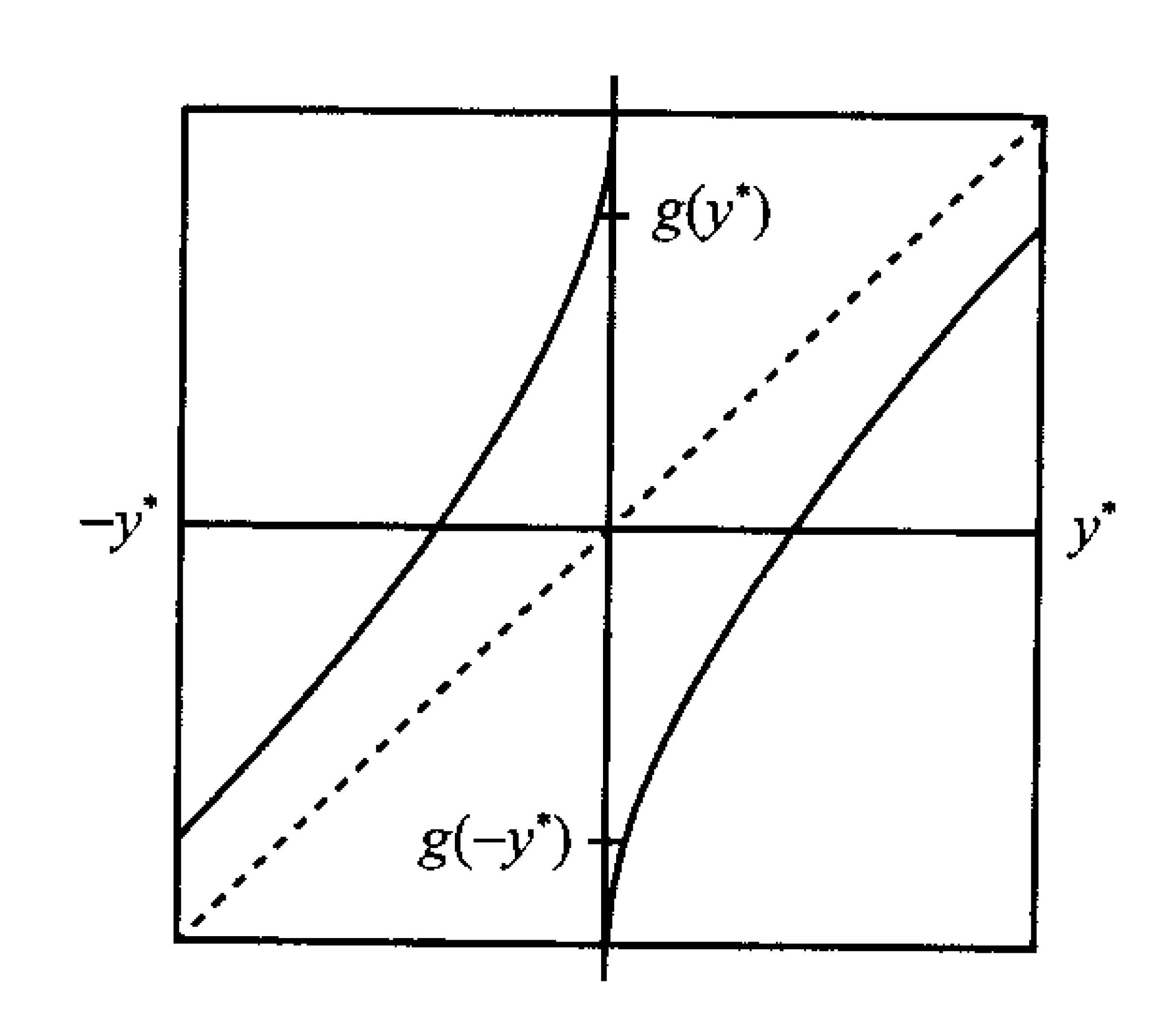


图 14.10　$I=[-y^*,y^*]$ 上的一维函数 g 的图像

令 $y_0\in I$. 我们来研究集合 $A\cap\{y=y_0\}$ 的结构. 我们定义 y_0 的(**前向**) **轨道**为集合 $(y_0,y_1,y_2,\cdots)$, 其中 $y_n=g(y_{n-1})=g^n(y_0)$. 对每个 y_0, 它的前向轨道是唯一确定的, 当然它有可能在 $g^n(y_0)=0$ 时停止.

y_0 的一个**后向轨道**是指一个形如 $(y_0,y_{-1},y_{-2},\cdots)$ 的序列, 其中 $g(y_{-k})=y_{-k+1}$. 与 g 的前向轨道不同, 对一个给定的 $y_0\neq\pm y^*$, 它有无穷多条不同的后向轨道 ($y_0=\pm y^*$ 在 I 中没有原像). 为了看出这点, 首先假设 y_0 不在 $\pm y^*$ 的正向轨道上. 从而, 由于每个 $y_{-k}\neq\pm y^*$, 它要么有一个原像要么有 2 个原像. 如果 y_{-k} 有一个原像 y_{-k-1}, 则 y_{-k} 位于 $(-y^*,g(-y^*))$ 或 $(g(y^*),y^*)$ 中. 但是, g 的图像表明 y_{-k-1} 必定有 2 个原像. 从而, 在一个给定的后向轨道中, 不可能有相邻的 2 个点都只有一个原像, 这就说明了 y_0 有无穷多条不同的后向轨道.

如果碰巧存在 $k>0$ 使得 $y_{-k}=\pm y^*$, 由于 $\pm y^*$ 在 I 中没有原像, 这条后向轨道必定停止. 但是, y_{-k+1} 必定有 2 个原像, 其中一个为 I 的一个端点, 另一个是不同于这个端点的 I 中的其他点. 于是, 我们可像前面一样继续取这第二个后向轨道的原像, 进而像前面一样生成无穷多条不同的后向轨道.

我们断言, y_0 的这些无穷多条后向轨道中的每条都对应于 $A\cap\{y=y_0\}$ 中唯一的一点. 为了看出这点, 考虑 R 中由 $y=y_{-k}$ 给出的线段 J_{-k}. 于是, 对每个 k, $\Phi^k(J_{-k})$ 都是 J_0 的一个闭子区间. 由于 $\Phi(y=y_{-k-1})$ 是 $y=y_{-k}$ 的一个真子区

间, 我们有 $\Phi(J_{-k-1}) \subset J_{-k}$. 于是区间套 $\Phi^k(J_{-k})$ 的交集非空, 而且根据构造, 这个交集中的每点都有后向轨道 $(y_0, y_{-1}, y_{-2}, \cdots)$. 进一步, 由于 Φ 的每次作用都将区间 $y = y_{-k}$ 的长度压缩一个因子 $c < 1$, 因而这个交集只包含唯一的一个点.

于是在洛伦茨系统的模型中, 我们就看到吸引子 $\mathcal{A}$ 是一个复杂的集合. 我们已经证明:

命题 洛伦茨系统模型中的吸引子 $\mathcal{A}$ 与 R 中每条线段 $y = y_0 \neq \pm y^*$ 都有无穷多个交点. 当时间向前时, 所有经过这条线段上每一点的解曲线要么

(1) 与线段 $y = 0$ 相交, 此时, 所有的解曲线都要趋于 $(0,0,0)$ 处的平衡点, 要么

(2) 连续不停地与 R 相交, 而且当时间增加时, 这些交点在线段 $y = y_k$ 上的距离趋于 0. ■

现在我们转到 Φ 在 R 中的动力学. 我们首先讨论一维函数 g 的行为, 然后利用这些信息来了解 Φ 的行为. 任给一点 $y_0 \in I$, 由于 $g' > \sqrt{2}$, y_0 附近点的前向轨道都要与 y_0 的前向轨道分开. 准确地说, 我们有:

命题 令 $0 < \nu < y^*$, $y_0 \in I = [-y^*, y^*]$. 任给 $\epsilon > 0$, 我们都可找到 $u_0, v_0 \in I$ 满足 $|u_0 - y_0| < \epsilon$, $|v_0 - y_0| < \epsilon$, 以及 $n > 0$ 满足 $|g^n(u_0) - g^n(v_0)| \geqslant 2\nu$.

证明 令 J 是中心在 y_0 长度为 2ϵ 的区间. g 的每次迭代都将 J 的长度扩张至少 $\sqrt{2}$ 倍, 从而存在一个迭代使得 $g^n(J)$ 的内部包含 0. 于是 $g^{n+1}(J)$ 就包含任意接近 $\pm y^*$ 的点, 从而这些 $g^{n+1}(J)$ 中的点彼此之间的距离就至少是 2ν. 这就完成了证明. ■

我们来解释这个命题在吸引子 A 中的意义. 任给这个吸引子中的一点, 我们总可以在它附近找到 2 个点, 使得它们的前向轨道分开到它们可能分开的那么远. 这是混沌系统的标志. 我们将这种行为称为**对初值条件的敏感依赖性**. 轨道的初始位置的一点微小变化将导致轨道最终行为的巨大改变. 注意, 对于在吸引子 $\mathcal{A}$ 中相应的流, 一定也有类似的敏感依赖性; 一些在 $\mathcal{A}$ 中邻近的解曲线也会分得很开. 这就是在图 14.2 中所看到的行为.

敏感依赖性应该与 A 中位于同一直线 $y =$ 常数 $(y \in (-y^*, y^*))$ 上的点的行为对照来看. 我们前面已经看到, A 中包含无穷多个这种点. 在 Φ 的迭代下, 所有这些点在逐次迭代下会越来越接近而不是分开.

回忆一下, I 中的一个子集是稠密的, 如果它的闭包是整个 I. 等价地有, I 中的一个子集是稠密的, 如果这个子集中有点任意接近 I 中任意给定的点. y_0 是 g 的一个**周期点**, 如果对某个 $n > 0$ 有 $g^n(y_0) = y_0$. g 的周期点对应于流的周期解.

命题 g 的周期解点在 I 中稠密.

证明 与上节证明 A 为吸引子一样, 任给 $I - \{0\}$ 的子区间 J, 我们可以找到 n, 使得 g^n 将某个子区间 $J' \subset J$ 一对一的映到 $(-y^*, 0]$ 或 $[0, y^*)$. 从而要么 $g^n(J')$ 包

含 J', 要么下一次迭代, $g^{n+1}(J')$ 包含 J'. 在任何一种情形, g^n 或 g^{n+1} 的图像都要穿过 J' 上的对角线 $y=x$. 这就得到了 g 在 J 中的一个周期点. ■

现在我们在吸引子 A 中来解释这个结果. 我们断言, Φ 的周期点也在 A 中稠密. 为了看出这点, 令 $P \in A$, U 为 P 的一个开邻域. 我们假设线段 $y=0$ 不穿过 U(否则, 只需取一个与 $y=0$ 不相交的更小的邻域就可). 对于充分小的 $\epsilon>0$, 我们构造一个矩形 $W \subset U$, 它的中心位于 P 处, x 方向的宽度为 2ϵ, y 方向的高度为 ϵ.

令 $W_1 \subset W$ 为中心在 P 处, 边长为 $\epsilon/2$ 的一个小正方形. 根据上节的传递性结果, 我们可以找到一点 $Q_1 \in W_1$ 使得 $\Phi^n(Q_1)=Q_2 \in W_1$. 如果需要的话取 W_1 的一个子集, 我们不妨假设 $n>4$, 而且进一步假设 n 大到满足 $c^n<\epsilon/8$. 于是就有 $\Phi^n(W)$(不是 $\Phi^n(W_1)$) 以近乎竖直的方式穿过 W 的内部, 而且延伸到它的上下边界之外, 如图 14.11 所示. 这个事实利用到了双曲性条件.

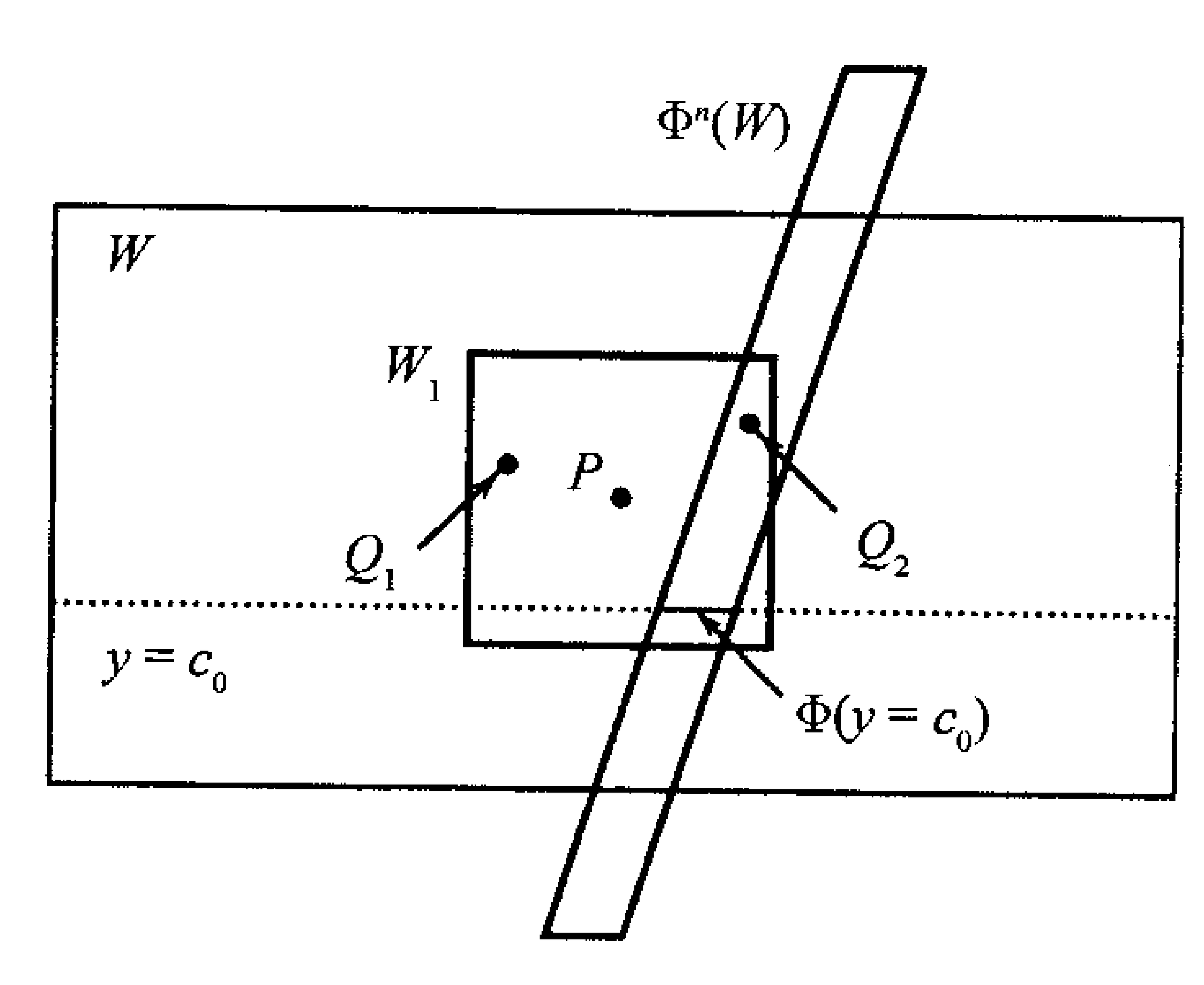


图 14.11 W 在 Φ^n 下的像穿过它自己

现在考虑 W 中的线段 $y=c$. 这些线段被 Φ^n 映到了 R 中其他这种线段. 由于竖直方向被扩张, 这些线段中某些必定被映到 W 的上方, 也有一些被映到 W 的下方. 从而, 必定有一条这种线段 $y=c_0$ 被 Φ^n 映到自己的内部, 因而在这条线段上必定有 Φ^n 的不动点. 由于这条线段包含在 W 中, 我们就得到了 Φ 在 W 中的一个周期点. 这就证明了 A 中周期点的稠密性.

对流来说, 从 Φ 的周期点出发的一个解就是一条闭轨. 位于闭轨上的点构成的集合是 $\mathcal{A}$ 的一个稠密子集. 从拓扑观点来看, 这些闭轨的结构是非常有趣的, 因为这些闭曲线中的许多都真正被“打结”了. 参见 [9] 和习题 10.

最后, 我们称一个函数 g 在 I 上是**传递的**, 如果对 I 中的任一对点 y_1 和 y_2 以及 y_i 的任一邻域 U_i, 我们都能找到 $\tilde{y} \in U_1$ 以及 n 使得 $g^n(\tilde{y}) \in U_2$. 与证明周期点的稠密性一样, 我们可以利用 g 在 y 方向为扩张的事实证明:

命题 函数 g 在 I 上是传递的. ■

我们将证明的细节留给读者. 对于 Φ 来说, 在证明 A 是一个吸引子时, 我们

就几乎已经证明了相应的结论. 我们唯一没有提供的细节就是如下事实：可以找到 A 中的一个点使得它的轨道可以任意接近 A 中任意给定的一对点. 对于这个细节, 参见习题 4.

于是, 我们可以将洛伦茨模型的 Φ 在吸引子 A 上的动力学总结如下：

定理 (**洛伦茨模型的动力学**) 洛伦茨模型限制在吸引子 A 上的庞加莱映射具有下面的性质：

(1) Φ 具有对初值条件的敏感依赖性;

(2) Φ 的周期点在 A 中稠密;

(3) Φ 在 A 上是传递的. ■

我们称具有上面三条性质的映射为**混沌的**. 我们提醒读者注意, 正如吸引子的定义一样, 混沌也有很多定义. 其中一些涉及轨道的指数分离性, 一些涉及**正李雅普诺夫指数**, 还有一些不要求周期点的稠密性. 有趣的是, 对于实线段上的连续函数, 周期点的稠密性和传递性就足以保证敏感依赖性. 见 [8]. 在下一章我们还将更深入地讨论离散系统的混沌行为.

14.6 探索：Rössler 吸引子

在这个探索中, 我们研究一个在很多方面都类似于洛伦茨系统的三维系统：Rössler 系统, 它由下面的方程给出：

$$
\begin{aligned}
x' &= -y - z, \\
y' &= x + ay, \\
z' &= b + z(x - c),
\end{aligned}
$$

其中 a, b, c 为实参数. 为了简单起见, 我们将注意力集中到如下的情形：$a = 1/4, b = 1$, 而 c 的变化范围为从 0 到 7.

与洛伦茨系统一样, 证明这个系统的一些特定结论是很困难的, 因而这个探索的大部分将以数值实验和构造模型为主.

(1) 首先找出这个系统的所有平衡点.

(2) 描述在 $c = 1$ 处发生的分岔.

(3) 用数值方法研究当 c 增加时, 这个系统的行为. 你观察到了什么样的分岔？

(4) 在图 14.12 中, 我们画出了当 $c = 5.5$ 时的一个解. 对这个参数值, 计算一些其他的解, 并从 $\mathbb{R}^3$ 的其他视点将计算结果示意出来. 对这个系统的行为, 你有什么样的猜测？

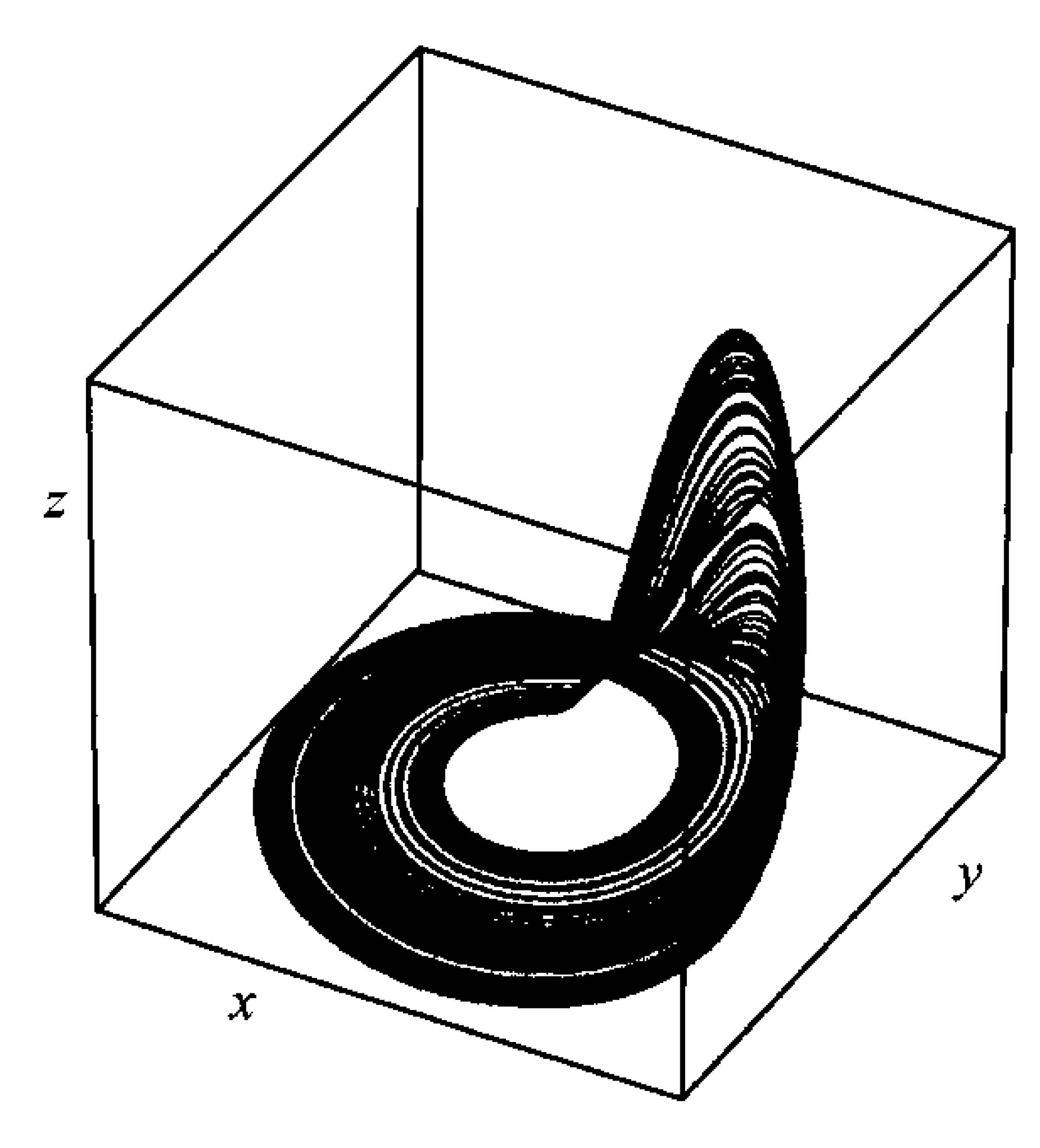


图 14.12 Rössler 吸引子

(5) 利用本章描述的技巧, 设计一个几何模型来模仿 Rössler 系统在这个参数值下的行为.

(6) 构造一个二维区域的模型映射, 使得它的动力学可以解释在这个系统中观察到的行为.

(7) 与洛伦茨系统中一样, 描述一个可能的方法将上述二维模型映射约化到一个区间上的映射.

(8) 给出这个一维模型映射的一个明确表达式. 你能对模型的混沌行为说些什么?

(9) 在 Rössler 系统中, 当 c 增加到大于 5.5 时, 你能观察到什么其他的分岔?

习　　题

1. 考虑系统

$$x' = -3x,$$
$$y' = 2y,$$
$$z' = -z.$$

从 14.4 节中可知, 存在函数 $h: \mathcal{R}_1 \to \mathcal{R}_2$, 其中 $\mathcal{R}_1$ 由 $|x| \leqslant 1, 0 < y \leqslant \epsilon < 1$ 和 $z = 1$ 给出, $\mathcal{R}_2$ 由 $|x| \leqslant 1, 0 < z \leqslant 1$ 和 $y = 1$ 给出. 证明 h 的表达式为:

$$h\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} xy^{3/2} \\ y^{1/2} \end{pmatrix}.$$

2. 交换上题中 x 和 z 的地位, 即, 假设 $x' = -x$ 且 $z' = -3z$. 在这种情形下, 描述 $h(x, y)$ 在 $\mathcal{R}_2$ 中的像.

3. 对于模型吸引子中的庞加莱映射 $\Phi(x,y)=(f(x,y),g(y))$, 利用习题 1 的结论证明当 $y\to 0$ 时, $g'(y)\to\infty$.
4. 证明在洛伦茨模型中, 可以用一个位于吸引子中的解验证传递性条件.
5. 证明在模型洛伦茨吸引子中任何点的任意近处, 都有一个解最终趋于位于 $(0,0,0)$ 处的平衡点.
6. 证明在模型洛伦茨系统中, 存在一个周期解 γ, 它与矩形 R 正好有两个不同的交点.
7. 证明在模型洛伦茨吸引子中任何点的任意近处, 都有一个解最终趋于上题中的周期解 γ.
8. 考虑如图 14.13 所示的矩形 R 上的映射 Φ, 其中 Φ 具有类似于模型洛伦茨吸引子中 Φ 的性质. Φ 有多少个周期为 n 的周期点?

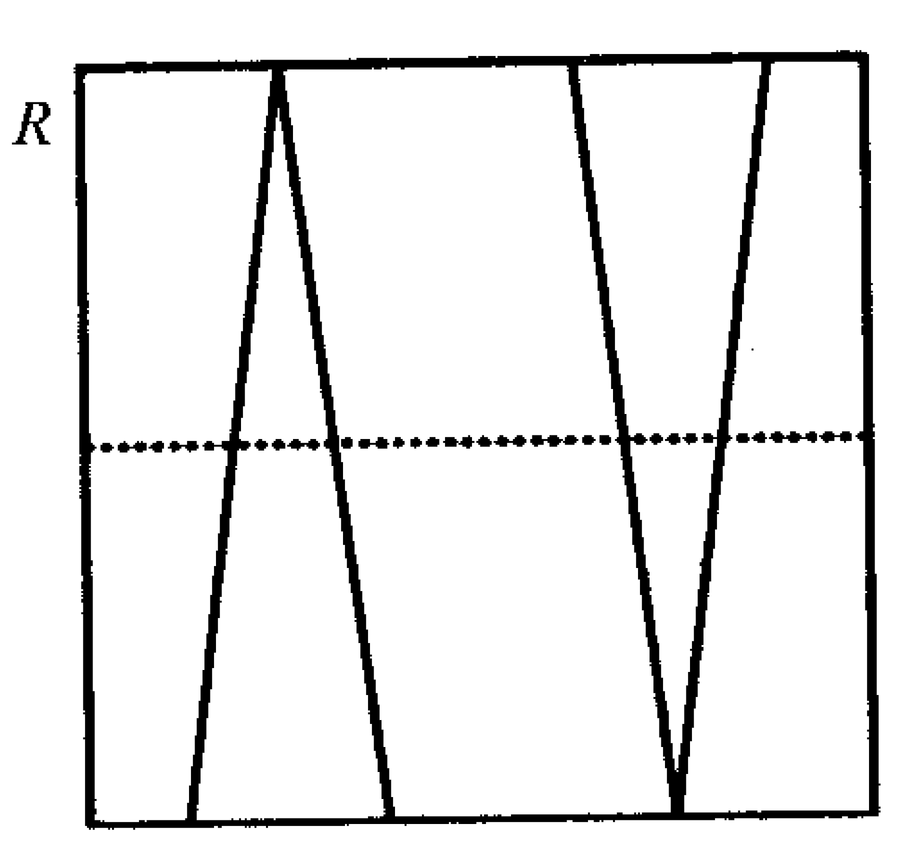


图14.13　R在Φ下的像完全穿过它自己

9. 考虑系统

$$\begin{aligned}x'&=10(y-x),\\ y'&=28x-y+xz,\\ z'&=xy-(8/3)z.\end{aligned}$$

证明这个系统在不是混沌的. (请注意 y' 的方程中 $+xz$ 这项.) 提示: 证明大多数解都趋于 ∞.

10. $\mathbb{R}^3$ 中的一条简单闭曲线是打了结的, 如果它不能连续地形变成 (在形变过程中不能出现自相交)xy 平面中"未打结的"单位圆周. 利用模型洛伦茨吸引子, 跟随 Φ 的动力学, 简略地画出一条打了结的曲线 (它应该近似于一个真实解). (为此你可能需要用到一根绳子!)
11. 利用计算机研究当 r 从 1 增加到 28 时洛伦茨系统的行为 $(\sigma=10, b=8/3)$. 定性地描述你所观察到的所有分岔.

第 15 章　离散动力系统

在本章我们的目的是开始研究离散动力系统. 本书中, 我们已经在多处看到, 有时候可以将微分方程的流的研究约化成一个函数 (即庞加莱映射) 迭代的研究. 这种约化具有几个优势. 首先, 也是最重要的, 庞加莱映射作用在一个维数较低的空间上, 从而这使得直观图像更加容易. 其次, 寻找离散系统的“解”, 我们不必积分, 而是对于一个给定的函数, 我们只是将它一次又一次地迭代, 就可确定这个轨道的行为, 这也就给出了相应的解的行为.

有了这两点简化后, 理解时常出现在微分方程系统中的复杂混沌行为就变得容易多了. 虽然离散动力系统的研究是一个可以轻易充塞这整本书的课题, 但是在本章我们将主要关注这一理论中有助于我们了解一维混沌行为的部分. 在随后的一章, 我们将这些想法推广到维数较高的情形.

15.1　离散动力系统介绍

在整个这一章, 我们将始终讨论实函数 $f:\mathbb{R}\to\mathbb{R}$. 我们照常始终假设 f 是 C^∞ 的, 但是有几个特殊例子并不满足这个假设.

记 f^n 为 f 的 n 次迭代. 即, f^n 是 f 与它自己的 n 重复合. 给定 $x_0\in\mathbb{R}$, x_0 的**轨道**是指序列

$$x_0,\ x_1=f(x_0),\ x_2=f^2(x_0),\cdots,\ x_n=f^n(x_0),\cdots.$$

点 x_0 称为这个轨道的**种子**.

例　令 $f(x)=x^2+1$, 则种子 0 的轨道为序列

$$\begin{aligned}
x_0 &= 0,\\
x_1 &= 1,\\
x_2 &= 2,\\
x_3 &= 5,\\
x_4 &= 26,\\
&\ \vdots\\
x_n &= \text{很大},\\
x_{n+1} &= \text{更大},\\
&\ \vdots
\end{aligned}$$

等等, 于是我们看到, 当 $n \to \infty$ 时, 这个轨道趋于 ∞. ■

与微分方程系统中的平衡解类似, 不动点在离散动力系统中非常重要. 点 x_0 称为一个**不动点**, 如果有 $f(x_0) = x_0$. 显然, 一个不动点的轨道就是常数序列: $x_0, x_0, x_0, \cdots$.

与微分方程中闭轨类似的是**周期为 n 的周期点**. 周期点是指对某个 $n > 0$ 满足 $f^n(x_0) = x_0$ 的种子 x_0. 从而周期轨道像闭轨一样复制它自己:

$$x_0, x_1, \cdots, x_{n-1}, x_0, x_1, \cdots, x_{n-1}, x_0, \cdots.$$

周期为 n 的周期轨道也称为一个**n 环**. 我们称周期点 x_0 具有**最小周期 n**, 如果 n 是满足 $f^n(x_0) = x_0$ 的最小正整数.

例 函数 $f(x) = x^3$ 的不动点位于 $x = 0, x = \pm 1$. 函数 $g(x) = -x^3$ 在 0 处有一个不动点; 由于 $g(1) = -1, g(-1) = 1$, 故 $g^2(\pm 1) = \pm 1$, 因而 g 在 $x = \pm 1$ 处有一个周期为 2 的周期点. 函数

$$h(x) = \frac{1}{2}(2 - x)(3x + 1)$$

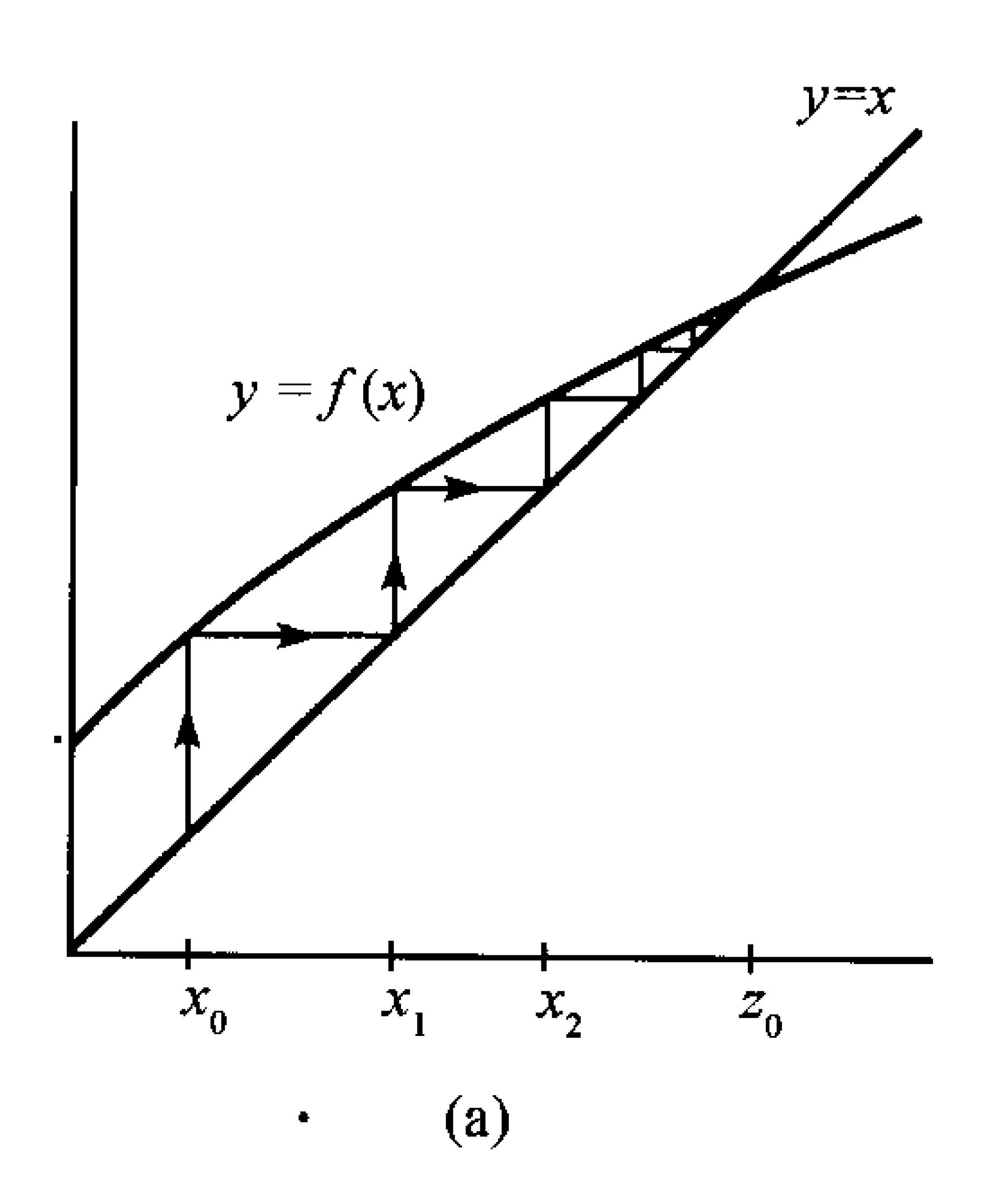
有一个 3 环, 它是 $x_0 = 0, x_1 = 1, x_2 = 2, x_3 = x_0 = 0, \cdots$. ■

直观地表示一维离散动力系统轨道的一个有用的方法是通过**图像迭代**. 在这个图中, 我们同时作出曲线 $y = f(x)$ 与对角线 $y = x$ 的图像. 我们如下显示 x_0 的轨道: 首先从对角线上的点 (x_0, x_0) 开始, 然后向着 f 的图像作一条竖直线段, 与之相交于 $(x_0, f(x_0)) = (x_0, x_1)$. 然后再作一水平线段回到对角线上的 (x_1, x_1). 这个过程使我们从对角线上正好在种子 x_0 上方的一点移动到了对角线上正好在 x_0 轨道的下一个点, 即 x_1. 接着, 我们从 (x_1, x_1) 继续: 先竖直走到 f 的图像上的点 (x_1, x_2), 然后水平地走回对角线上的 (x_2, x_2). 从 x 轴上看, 这使我们从 x_1 移动到了 x_0 轨道上的下一个点, 即 x_2. 继续进行下去, 我们就产生了一对线段的序列, 其中每对线段都终止在对角线上一个形如 (x_n, x_n) 的点.

在图 15.1a 中, 图像迭代显示, 在 f 的迭代下, x_0 的轨道趋于不动点 z_0. 在图 15.1b 中, 图像迭代则显示 x_0 在 g 下的轨道位于一个 3 环上: $x_0, x_1, x_2, x_0, x_1, \cdots$.

与微分方程的平衡点一样, 对于离散动力系统, 不动点有各种不同的类型. 假设 x_0 是 f 的一个不动点. 我们称 x_0 是 f 的一个**汇点**或一个**吸引不动点**, 如果存在 x_0 在 $\mathbb{R}$ 中的一个邻域 $\mathcal{U}$, 使得对所有的 $y_0 \in \mathcal{U}$ 和所有的 n 都有 $f^n(x_0) \in \mathcal{U}$, 并且当 $n \to \infty$ 时, $f^n(y_0) \to x_0$. 类似地, x_0 称为一个**源点**, 或一个**排斥不动点**, 如果 $\mathcal{U}$ 中的所有的点 (x_0 除外) 在 f 迭代下的轨道都要离开 $\mathcal{U}$. 一个不动点称为是**中性的**, 或**中立的**, 如果它既不是吸引的也不是排斥的.

我们已经看到, 对于微分方程, 正是向量场在一个平衡点处的导数决定了该平衡点的类型. 对于不动点情形也是如此, 只不过数字上发生了一点变化.



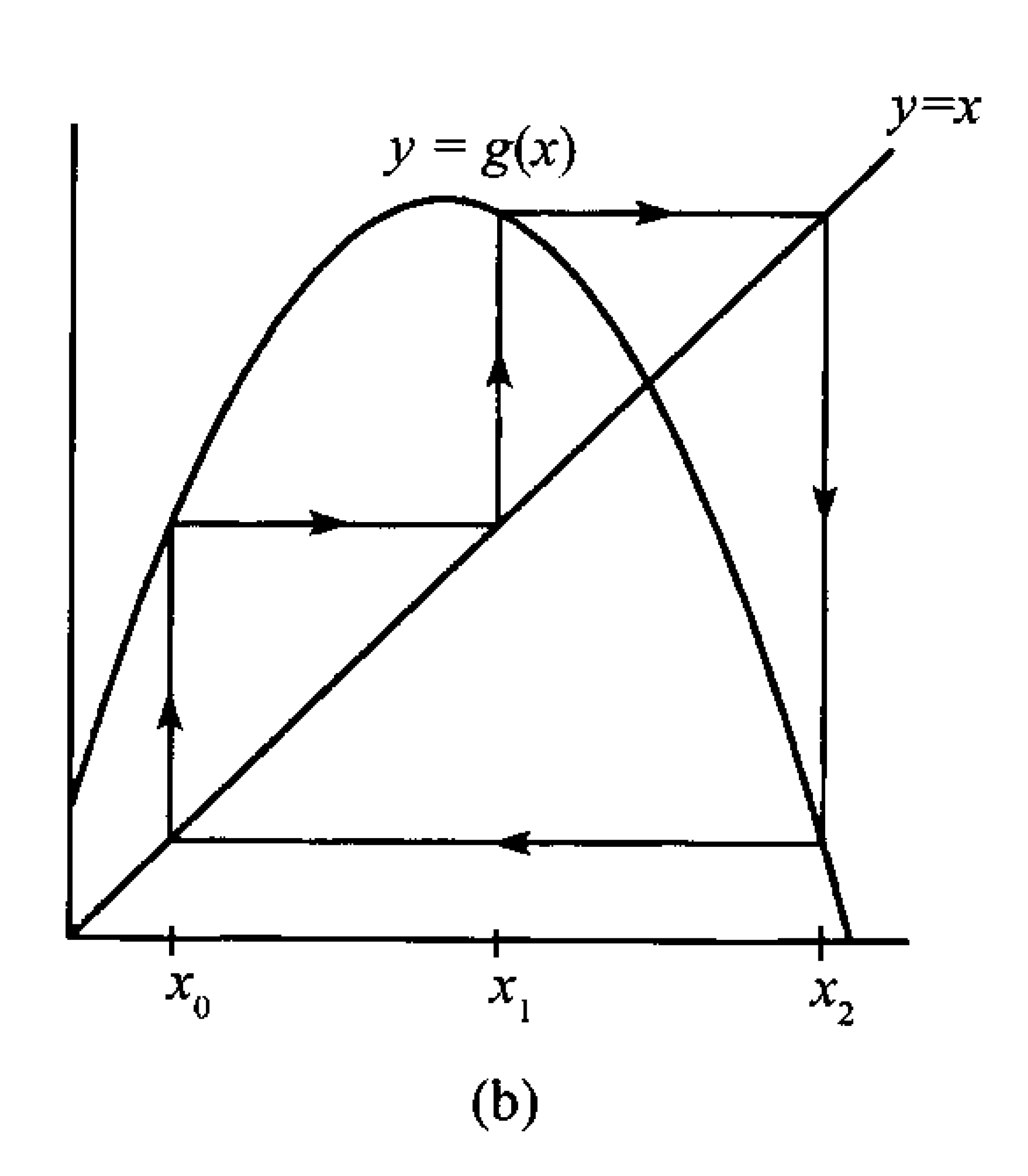


图 15.1 (a) 在 f 的迭代下, x_0 的轨道趋于不动点 z_0, 在 (b) 中, x_0 在 g 迭代下的轨道位于一个 3 环上

命题 假设 f 在 x_0 处有一个不动点 x_0.

(1) 如果 $|f'(x_0)|<1$, 则 x_0 是一个汇点;

(2) 如果 $|f'(x_0)|>1$, 则 x_0 是一个源点;

(3) 如果 $f'(x_0)=\pm 1$, 我们无法由此得出平衡点类型的任何信息.

证明 我们首先证明情形 (1). 假设 $|f'(x_0)|=\nu<1$. 选取 K 满足 $\nu<K<1$. 由于 f' 连续, 我们可以找到 $\delta>0$, 使得对于所有的 $x\in I=[x_0-\delta,x_0+\delta]$, 都有 $|f'(x)|<K$. 根据微分中值定理, 对任何 $x\in I$, 我们都有

$$\frac{f(x)-x_0}{x-x_0}=\frac{f(x)-f(x_0)}{x-x_0}=f'(c),$$

其中 c 为介于 x 和 x_0 之间的某个数. 于是我们就有

$$|f(x)-x_0|<K|x-x_0|.$$

由此可知, $f(x)$ 比 x 更靠近 x_0, 因而 $f(x)\in I$. 再次利用这个结果可得

$$|f^2(x)-x_0|<K|f(x)-x_0|<K^2|x-x_0|,$$

如此继续下去, 我们就有

$$|f^n(x)-x_0|<K^n|x-x_0|,$$

由于 $0<K<1$, 因而对所有的 $x\in I$, 都有 $f^n(x)\to x_0$.

情形 (2) 的证明是类似的. 对于情形 (3), 我们来看下面的 3 个函数:

(1) $f(x)=x+x^3$;

(2) $g(x) = x - x^3$;

(3) $h(x) = x + x^2$.

可以看出每个都以 0 点为不动点, 并且都有 $f'(0) = 1$. 但是图像迭代 (图 15.2) 表明: 0 是 f 的源点, 同时是 g 的汇点, 而对于函数 h 来说, 0 吸引一边, 排斥另一边. ■

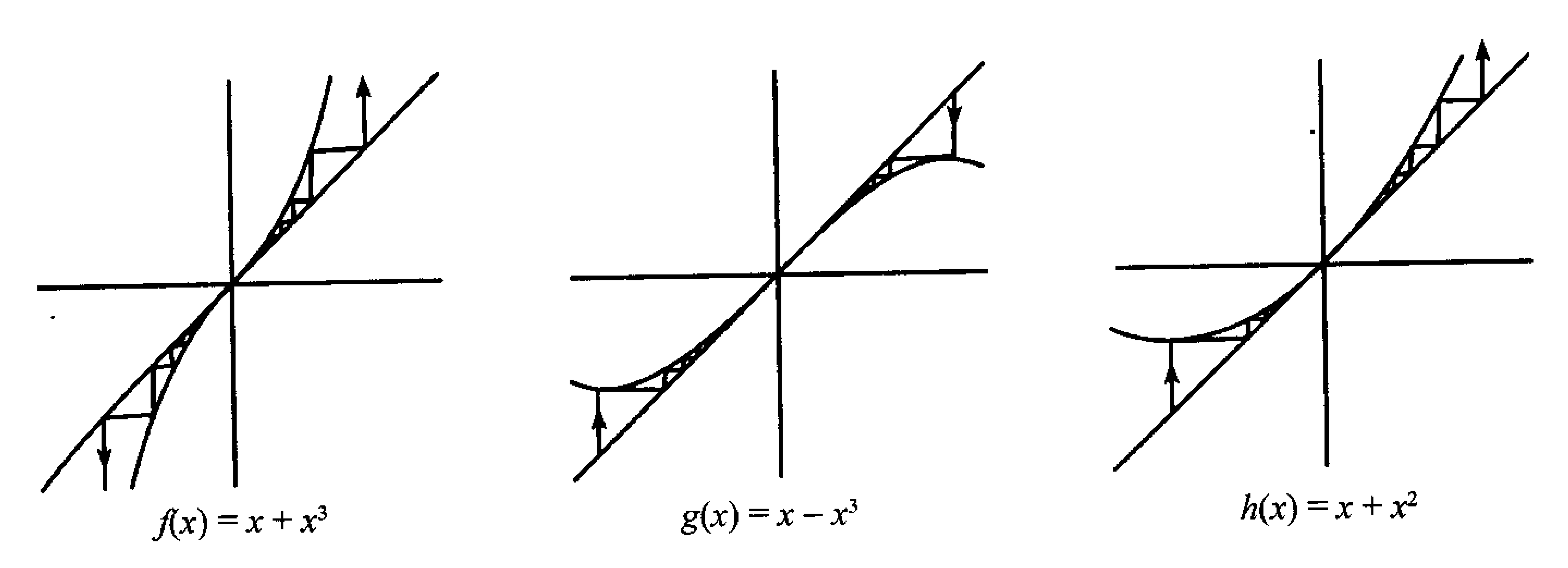


图 15.2 在每种情形, 0 点的导数都是 1, 0 是 f 的源点和 g 的汇点, 但既不是 h 的源点也不是 h 的汇点

注意, 如果在不动点 x_0 处有 $f'(x_0) < 0$, 在每次迭代下, x_0 附近点的轨道都从 x_0 的一边跳到另一边 (见图 15.3). 从中可以看出, 为什么常常称这种图像迭代的产物为一个**蛛网图**.

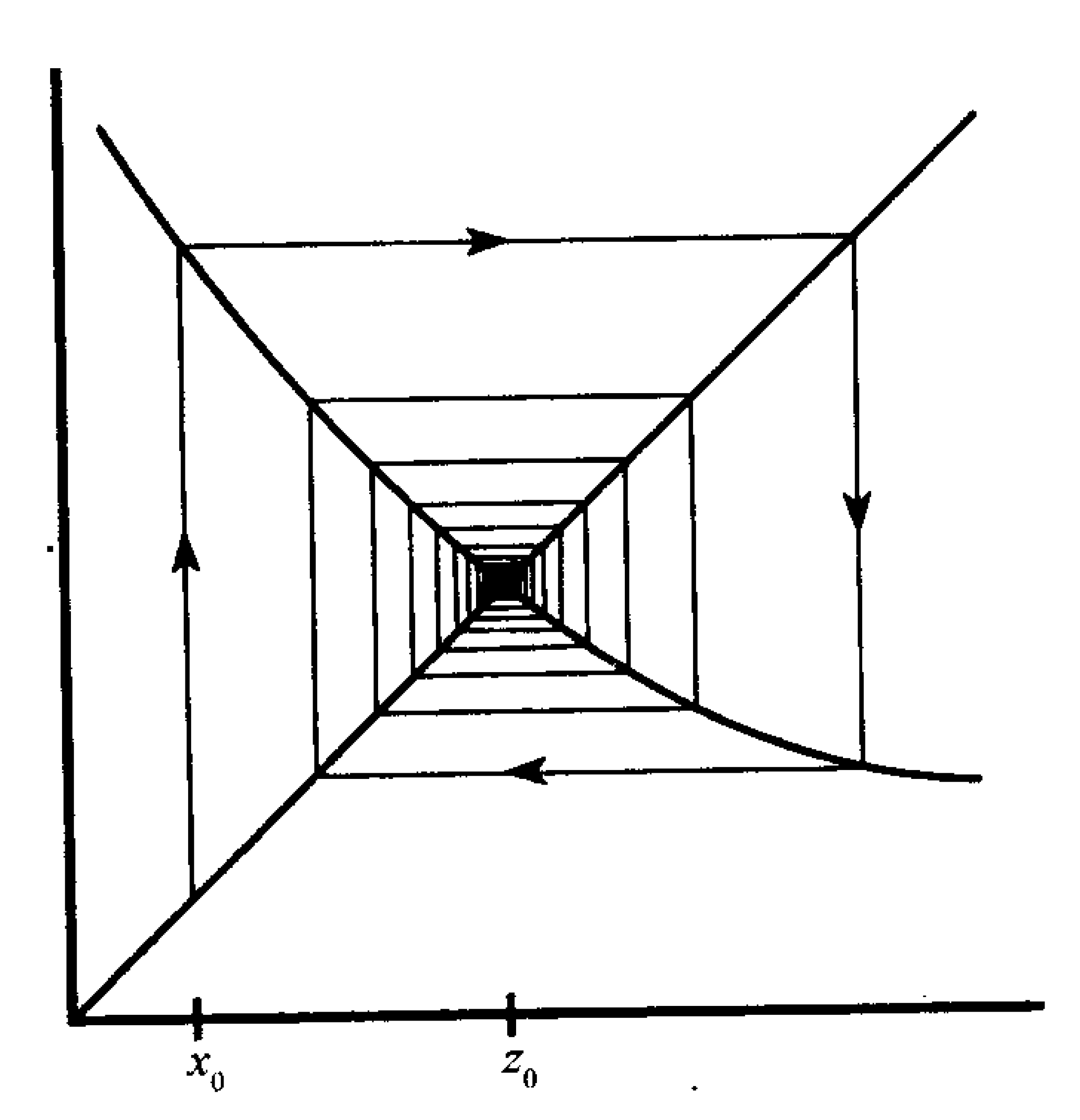


图 15.3 由于 $-1 < f'(z_0) < 0$, x_0 的轨道 "盘旋地" 趋于 z_0 处的不动点

由于 f 的一个周期为 n 的周期点 x_0 是 f^n 的一个不动点, 我们可以根据 $|(f^n)'(x_0)| < 1$, 或 $|(f^n)'(x_0)| > 1$ 将它们分别称为**汇点或源点**. 可以验证, 对于周期轨道上的任一点 x_j 都有, $(f^n)'(x_0) = (f^n)'(x_j)$, 因而这个定义是有意义的 (见本

章的习题 6).

例 函数 $f(x) = x^2 - 1$ 有一个由 $0, -1$ 构成的 2 环. 容易验证, $(f^2)'(0) = 0 = (f^2)'(-1)$, 因而这个环是一个汇. 在图 15.4 中, 我们同时作出了 f 的图像迭代和 f^2 的图像. 注意, 0 和 1 都是 f^2 的吸引不动点.

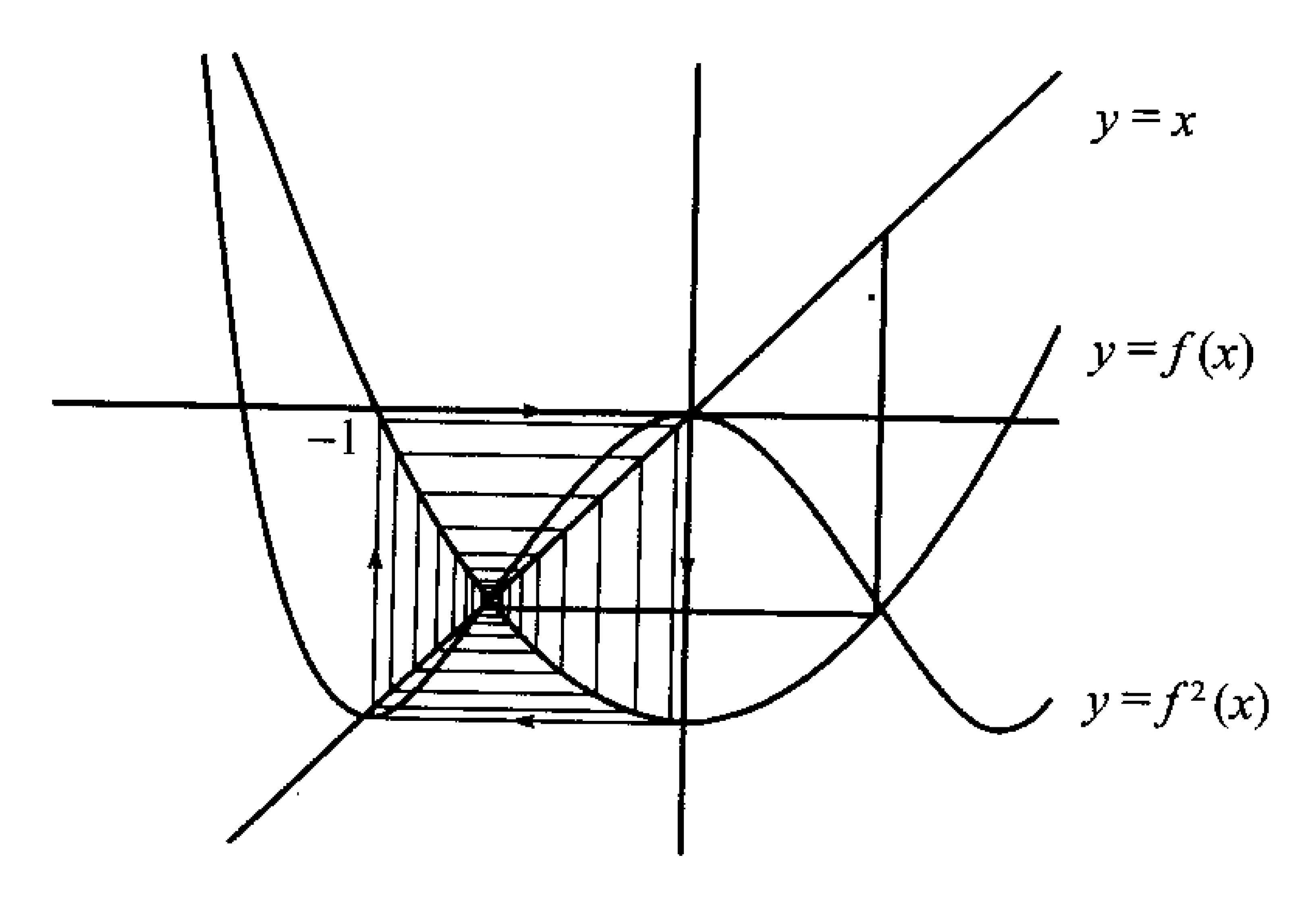


图 15.4 $f(x) = x^2 - 1$ 的图像以及 f^2 的图像表明 $0, -1$ 位于 f 的一个吸引 2 环上

15.2 分 岔

与微分方程一样, 当参数变化时, 离散动力系统也会发生分岔. 在本节我们讨论一维系统中出现的几类分岔.

例 令 $f_c(x) = x^2 + c$, 其中 c 是一个参数. 这族函数的不动点可通过求解方程 $x^2 + c = x$ 得到, 它们是

$$p_{\pm} = \frac{1}{2} \pm \frac{\sqrt{1-4c}}{2}.$$

于是, 当 $c > 1/4$ 时, 没有不动点; 当 $c = 1/4$ 时, 有唯一的不动点, 它位于 $x = 1/2$; 当 $c < 1/4$ 时, 有一对不动点 $p_{\pm}$. 图像迭代表明: 当 $c > 1/4$ 时, f_c 的所有轨道都趋于 ∞; 当 $c = 1/4$ 时, 位于 $x = 1/2$ 的不动点是中性的 (见图 15.5). 而当 $c < 1/4$ 时, 我们有 $f_c'(p_+) = 1 + \sqrt{1-4c} > 1$, 于是 p_+ 总是排斥的. 直接的计算还表明, 当 $-3/4 < c < 1/4$ 时, $-1 < f_c'(p_-) < 1$. 于是当 c 取这些值时, p_- 是吸引的. 当 $-3/4 < c < 1/4$ 时, 区间 $(-p_+, p_+)$ 中的所有轨道都趋于 p_-(但是, 严格来讲, 由于 $-p_-$ 被直接映到了 p_-, 因而它最终是不动的, 当 $c < 0$ 时, 这个区间中的某些其他轨道也是如此). 这样, 当 c 下降地通过分岔值 $c = 1/4$ 时, 我们看到首先出现单个的中性不动点, 然后立刻分裂成两个不动点, 其中一个是吸引的, 而另一个是排斥的. 这是一个**鞍结分岔**, 也称为**切分岔**. 图像上, 这种分岔本质上与第 8 章中对一阶微分方程描述的同名分岔是一样的 (见图 15.5). ■

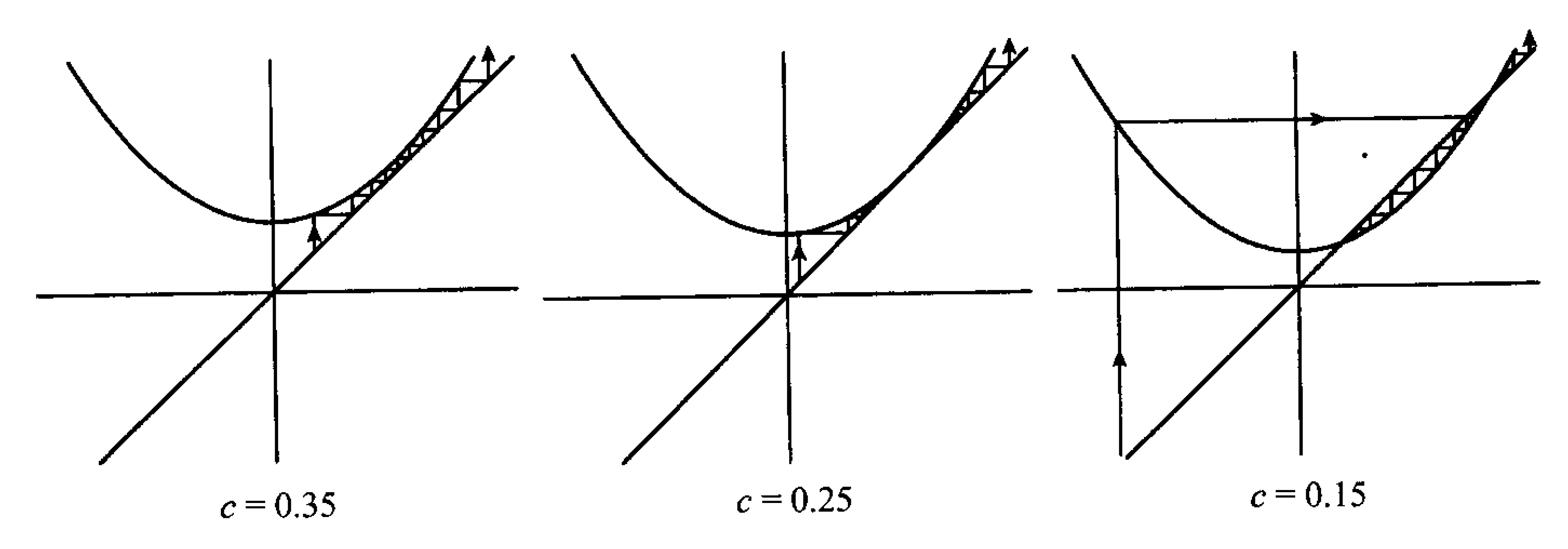


图 15.5 $f_c(x)=x^2+c$ 在 $c=1/4$ 时的鞍结分岔

注意, 在这个例子中, 在分岔值处, 不动点处的导数为 1. 这并不是偶然的, 因为我们有:

定理 (**分岔判据**) 令 f_λ 是一族连续依赖于参数 λ 的函数. 假设 $f_{\lambda_0}(x_0)=x_0$, 并且 $f'_{\lambda_0}(x_0)\neq 1$. 则存在关于 x_0 的一个区间 I, 一个关于 λ_0 的区间 J, 以及一个光滑函数 $p:J\to I$ 使得 $p(\lambda_0)=x_0$, 并且 $f_\lambda(p(\lambda))=p(\lambda)$. 进一步, f_λ 在 I 中没有其他的不动点.

证明 考虑由 $G(x,\lambda)=f_\lambda(x)-x$ 定义的函数. 根据假设, 有 $G(x_0,\lambda_0)=0$, 并且

$$\frac{\partial G}{\partial x}(x_0,\lambda_0)=f'_{\lambda_0}(x_0)-1\neq 0.$$

由隐函数定理知, 存在关于 x_0 的一个区间 I, 一个关于 λ_0 的区间 J, 以及一个光滑函数 $p:J\to I$ 使得对所有的 $\lambda\in J$, $p(\lambda_0)=x_0$, 并且 $G(p(\lambda),\lambda)=0$. 进一步, 除了 $x=p(\lambda)$ 外, $G(x,\lambda)\neq 0$. 证毕. ■

作为这个结果的一个推论, 只有在 f_λ 有一个导数为 1 的不动点时, f_λ 才会出现一个与不动点个数变化有关的分岔. 在这种参数值处出现的典型分岔就是鞍结分岔 (见习题 18 和习题 19). 但是, 还会出现许多其他类型的分岔.

例 令 $f_\lambda(x)=\lambda x(1-x)$. 注意, 对所有的 λ 都有 $f_\lambda(0)=0$. 我们有 $f'_\lambda(0)=\lambda$, 因而在 $\lambda=1$ 处我们可能会有一个分岔. f_λ 在 $x_\lambda=(\lambda-1)/\lambda$ 还有第二个不动点. 当 $\lambda<1$ 时, x_λ 是负的, 而当 $\lambda>1$ 时, x_λ 是正的. 当 $\lambda=1$ 时, x_λ 与 0 处的不动点重合, 从而 f_1 只有一个不动点. 计算表明, 当 $\lambda>1$(并且 $\lambda<3$) 时, 0 是排斥的而 x_λ 是吸引的, 而当 $\lambda<1$ 时, 情况正好相反. 因为这个原因, 这种分岔被称为**交换分岔**. ■

例 考虑函数族 $f_\mu(x)=\mu x+x^3$. 当 $\mu=1$ 时, 我们有 $f_1(0)=0$, 并且 $f'_1(0)=1$, 从而有可能出现分岔. 不动点为 0 和 $\pm\sqrt{1-\mu}$, 从而当 $\mu<1$ 时我们有 3 个不动点, 而当 $\mu\geqslant 1$ 时, 只有一个不动点. 因而, 当 μ 经过 1 时, 的确出现了分岔. ■

对于一维离散动力系统, 出现分岔的其他唯一情形是不动 (周期) 点处的导数取 -1, 此时, 不动点有可能从汇点变成源点, 或从源点变成汇点. 当不动点处的导

数取所有其他值时, 在参数发生微小变化时, 汇点依然为汇点, 源点依然为源点, 而且附近没有其他周期轨道. 周期轨道的某一部分可能接近源点, 但是不可能整条轨道都在源点附近 (见习题 7). 在不动点的导数取 -1 时, 出现的典型分岔是**倍周期分岔**.

例 作为这种类型分岔的一个简单例子, 我们来考虑函数族 $f_\lambda(x) = \lambda x$(在 $\lambda_0 = -1$ 附近考虑). 对所有的 λ, 0 都是一个不动点. 当 $-1 < \lambda < 1$ 时, 0 是一个吸引不动点, 所有其他的轨道都趋于 0. 当 $|\lambda| > 1$ 时, 0 是排斥的, 而所有的非零解都趋于 $\pm\infty$. 当 $\lambda = -1$ 时, 0 是一个中性不动点, 而所有非零解都位于 2 环上. 当 λ 穿过 -1 时, 不动点的类型从吸引变为排斥, 同时出现了一族 2 环. ■

一般地, 当出现倍周期分岔时, 不仅仅只对一个参数值才会存在 2 环. 这种分岔的一个更加典型的例子如下.

例 还是考虑 $f_c(x) = x^2 + c$, 这次关心的是 $c = -3/4$ 附近的 c. 我们在

$$p_- = \frac{1}{2} - \frac{\sqrt{1-4c}}{2}$$

处有一个不动点. 我们已经看到 $f'_{-3/4}(p_-) = -1$, 并且当 c 比 $-3/4$ 稍大时, p_- 是吸引的, 而当 c 比 $-3/4$ 稍小时, p_- 是排斥的. 图像迭代表明, 当 c 下降地经过 $-3/4$ 时, 产生了一个吸引的 2 环. 这就是倍周期分岔 (见图 15.6). 事实上, 可以很容易解出这些周期为 2 的周期点并验证它们就是吸引的 (当 $-5/4 < c < -3/4$ 时; 见习题 8). ■

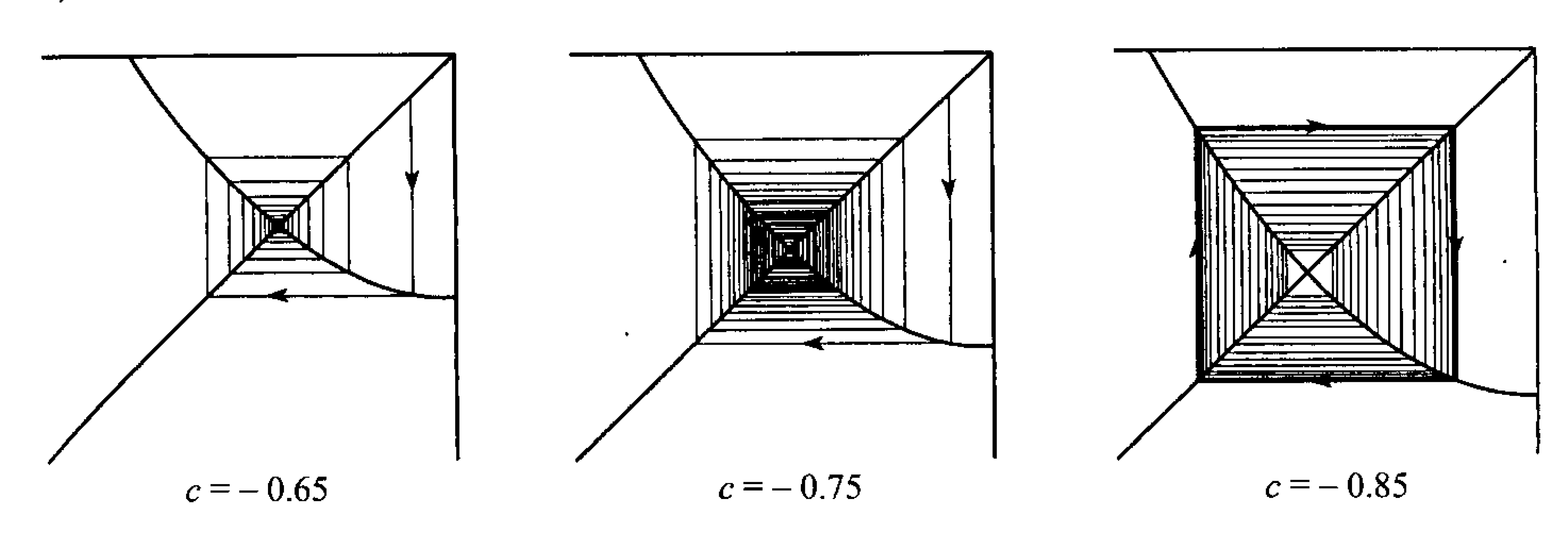


图 15.6 $f_c(x) = x^2 + c$ 在 $c = -3/4$ 时的倍周期分岔. 当 $c \geqslant -3/4$ 时, 不动点是吸引的, 而当 $c < -3/4$ 时, 它是排斥的

15.3 离散的合理模型

在第 1 章中, 我们介绍了一个最简单的一阶非线性微分方程, 即总量增长的合理模型

$$x' = ax(1-x).$$

在这个模型中, 我们考虑到了, 对一种典型的物种, 都存在一个最高容量, 而且我们还看到, 得出的解的行为十分简单：所有的非零解都趋于一个“理想的”物种总量. 但物种的总量并不是时间的连续函数！这就有必要重新考察这个模型. 一个更自然的方法是在一些特定的时刻, 比如说每年或每一代, 来统计物种的总量. 我们这里要介绍的就是这样一种模型：物种总量增长的**离散合理模型**.

假设我们每年 (或在其他特定时刻) 统计一次某种物种的总量. 记 x_n 为第 n 年年底的总量. 如果我们假设不会出现过度拥挤的现象, 则这样的物种总量模型是指数增长的, 我们假设

$$x_{n+1} = kx_n,$$

其中 $k > 0$ 是某个常数. 即, 下一年的总量正比于当年的总量, 于是我们就有

$$\begin{aligned} x_1 &= kx_0, \\ x_2 &= kx_1 = k^2x_0, \\ x_3 &= kx_2 = k^3x_0, \\ &\vdots \end{aligned}$$

显然, $x_n = k^n x_0$, 于是我们可以得出：当 $k > 1$ 时, 总量会爆炸, 而当 $0 \leqslant k < 1$ 时, 这种物种会灭绝, 或者当 $k = 1$ 时, 总量始终是一个常数.

这是一个一阶**差分方程**的例子, 即通过 x_{n-1} 的值就可以确定 x_n. 而二阶差分方程就是通过 x_{n-1} 和 x_{n-2} 的值来确定 x_n 的值. 从我们的观点来看, 上面指数增长模型中的相继总量就是通过简单地迭代函数 $f_k(x) = kx$(种子为 x_0) 得到的.

关于物种总量增长的一个更实际的假设是, 存在一个最大的物种总量 M, 也就是说, 如果总量超过了这个数量, 则所有的资源都被用尽了, 因而下一年物种将灭绝.

离散的**合理物种总量模型**就是一种反映了这些假设的模型. 这里我们假设总量的增长遵循规则

$$x_{n+1} = kx_n\left(1 - \frac{x_n}{M}\right),$$

其中 k, M 都是正的参数. 注意, 如果 $x_n \geqslant M$, 则 $x_{n+1} \leqslant 0$, 即在接下来的一年, 物种将灭绝.

我们并不直接处理实际的物种总量, 而是用 x_n 记总量占最大总量的比例, 从而 $0 \leqslant x_n \leqslant 1$. 于是, 合理的差分方程就变成了

$$x_{n+1} = \lambda x_n(1 - x_n),$$

其中 $\lambda > 0$ 是一个参数. 于是, 通过简单地迭代二次函数 $f_\lambda(x) = \lambda x(1 - x)$ (也称为**合理映射**), 我们就可以预测一个初始总量为 x_0 的物种的命运. 听起来很容易,

是不是? 但是直到 20 世纪 90 年代后期, 由于几百位数学家的努力, 这个简单的二次函数迭代才被完全地了解. 随后, 我们将看到, 为什么离散的合理模型要比它的亲戚 —— 合理的微分方程, 复杂这么多, 但是我们首先看一些简单情形.

我们只在单位区间 $I=[0,1]$ 上考虑合理映射. 我们有 $f_\lambda(0)=0$, 因而 0 总是一个不动点. 当 $0<\lambda\leqslant 1$ 时, 这个不动点在 I 中是吸引的, 之后, 它就成为排斥的. 由于 $f_\lambda(1)=0$, 故 1 这个点最终是不动的. 当 $\lambda>1$ 时, 在 I 中还有第二个不动点 $x_\lambda=(\lambda-1)/\lambda$. 不动点 x_λ 在 $1<\lambda\leqslant 3$ 时是吸引的, 而在 $\lambda>3$ 时是排斥的. 在 $\lambda=3$ 处, 出现了一个倍周期分岔 (见习题 4). 对于 3 和差不多 3.4 之间的 λ 的值, 唯一出现的周期点就是两个不动点和上面倍周期分岔出现的 2 环.

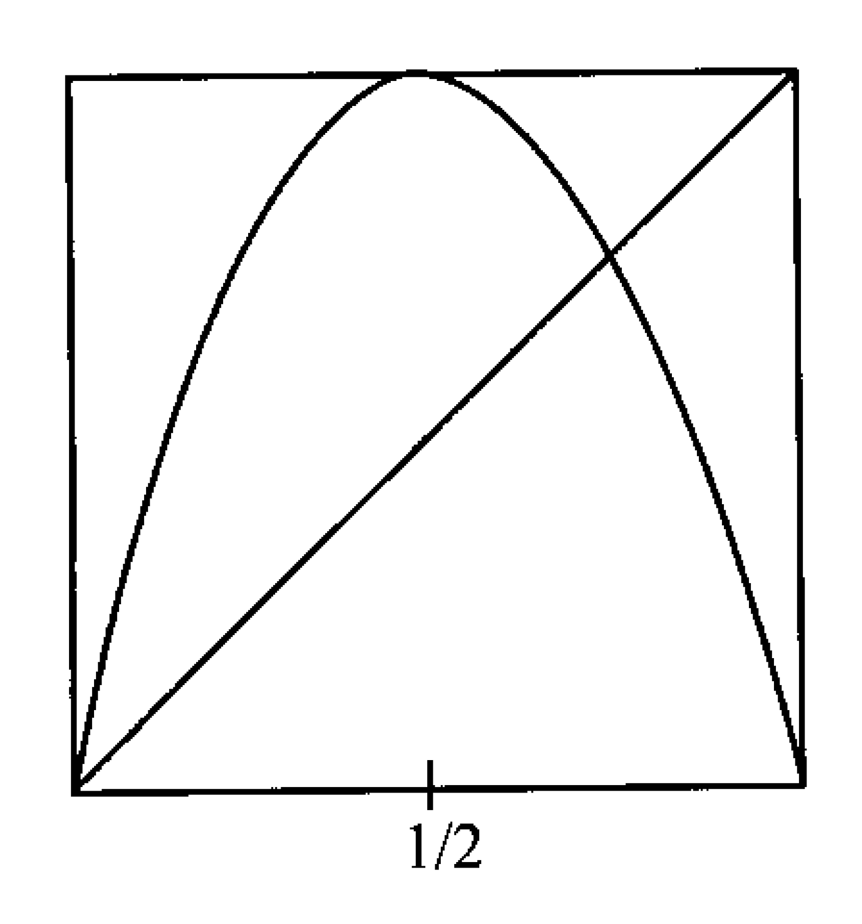
当 $\lambda=4$ 时, 情况会变得非常复杂. 注意 $f_\lambda'(1/2)=0$, 并且对于所有的 λ, 1/2 都是唯一的临界点. 当 $\lambda=4$ 时, 我们有 $f_4(1/2)=1$, 从而 $f_4^2(1/2)=0$. 于是 f_4 将每个半区间 $[0,1/2]$ 和 $[1/2,1]$ 都映满整个区间 I. 因而, 存在两个点 $y_0\in[0,1/2]$, $y_1\in[1/2,1]$ 使得 $f_4(y_j)=1/2$, 从而 $f_4^2(y_j)=1$. 这样就有

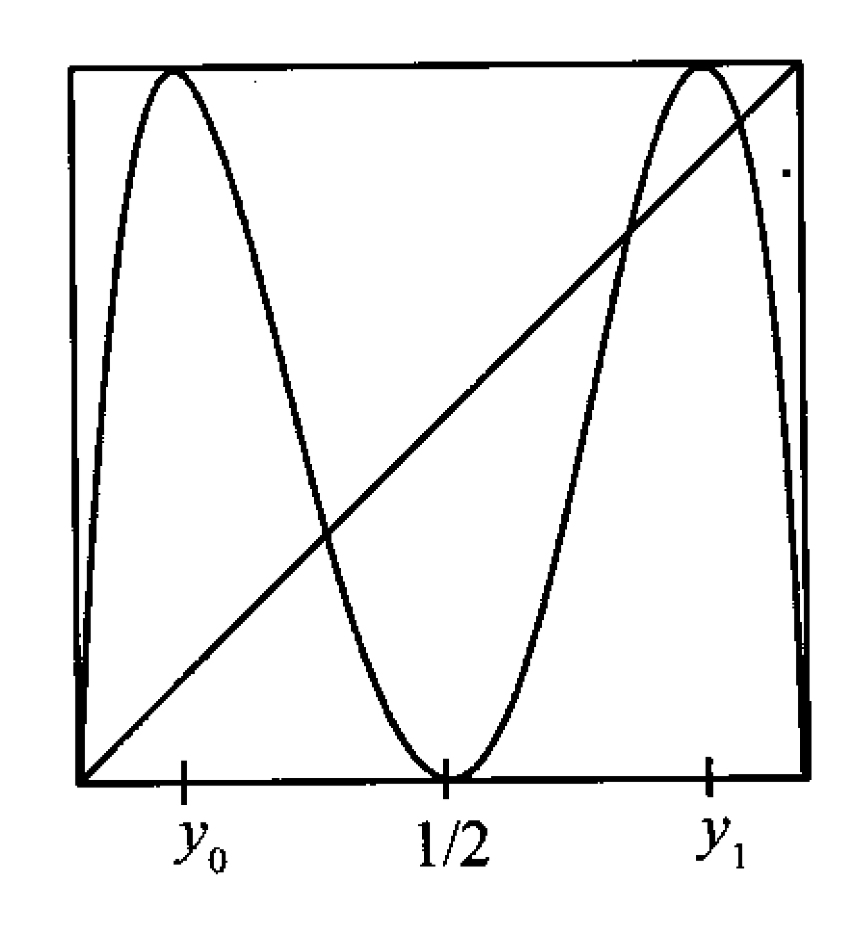
$$f_4^2[0,y_0]=f_4^2[y_0,1/2]=I$$

以及

$$f_4^2[1/2,y_1]=f_4^2[y_1,1]=I.$$

由于函数 f_4^2 是四次的, 故它的图像如图 15.7 所示. 按这种方式继续下去, 我们将找到 I 的 2^3 个子区间, 它们在 f_4^3 下都映满 I, 以及 2^4 个子区间, 它们在 f_4^4 下都映满 I, 等等. 于是我们就看到, f_4 在 I 中有 2 个不动点; f_4^2 在 I 中有 4 个不动点; f_4^3 在 I 中有 2^3 个不动点; f_4^4 在 I 中有 2^4 个不动点; 然后, 归纳地有, f_4^n 在 I 中有 2^n 个不动点. f_4 的 2 个不动点出现在 0 和 3/4. f_4^2 的 4 个不动点, 包括这 2 个不动点和一对周期为 2 的周期点. 在 f_4^3 的 8 个不动点中, 2 个是 f_4 的不动点, 而其他 6 个位于 f_4 的一对 3 环中. 在 f_4^4 的 16 个不动点中, 2 个是 f_4 的不动点, 2 个是周期为 2 的周期点, 剩下的 12 个都是周期为 4 的周期点. 显然, 当 λ 从 3.4 变化到 4 时, 发生了许多变化.





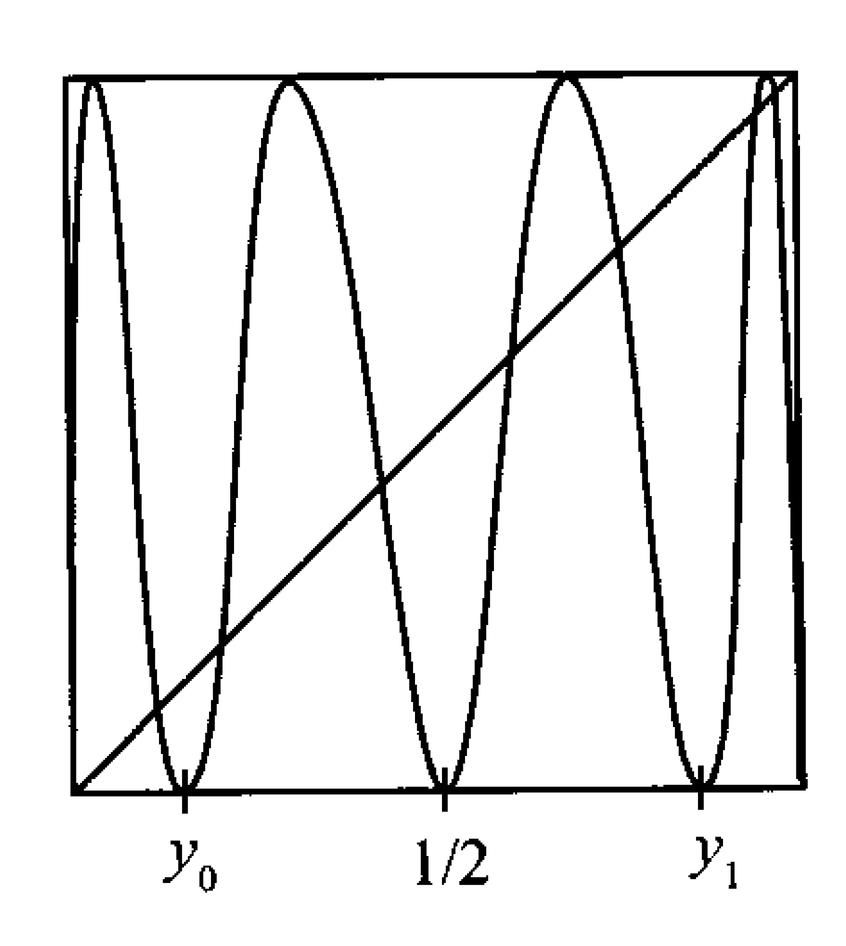
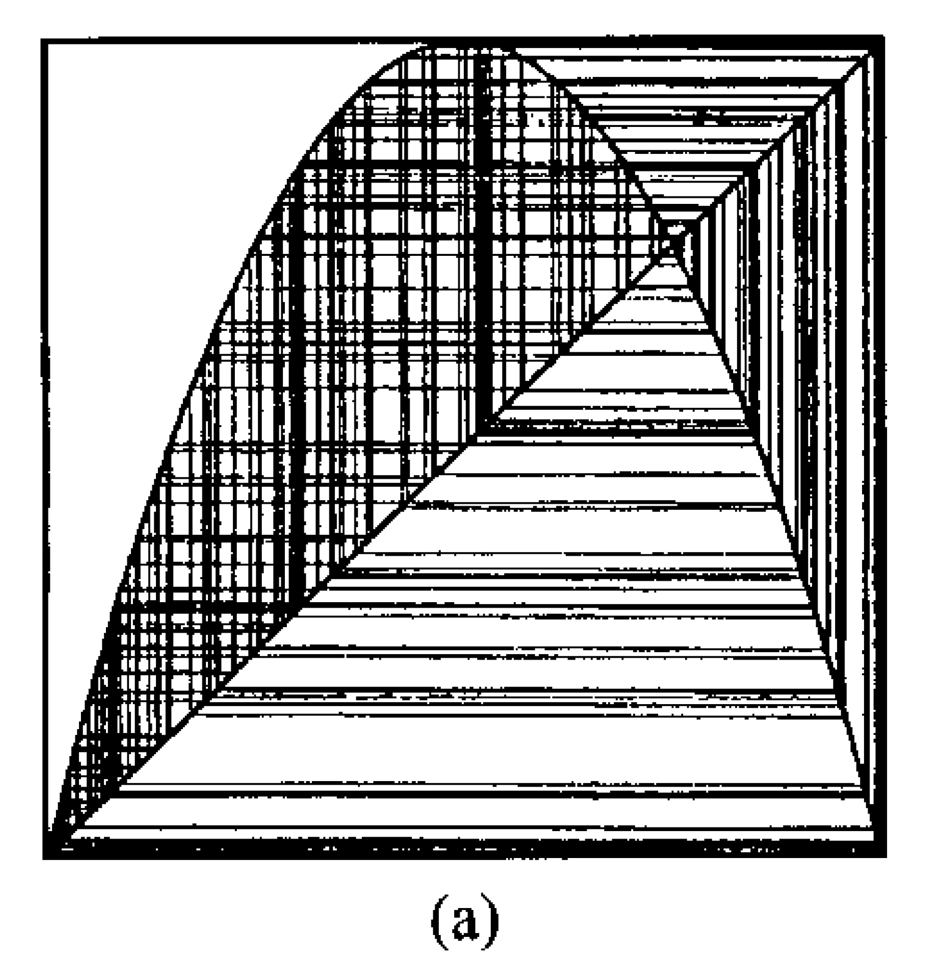


图 15.7 合理映射 $f_4=4x(1-x)$, 以及 f_4^2 和 f_4^3 在 I 上的图像

另一方面，如果我们在区间 I 中随机地选取一个种子，然后利用图像迭代作出它在 f_4 的迭代下的轨道，我们将很难看到这种环. 在图 15.8 中，我们画出了种子 0.123 在函数 f_4 迭代下的轨道，其中，分别取了 200 次迭代和 500 次迭代. 那里似乎发生了一些“混沌”现象.


(a)

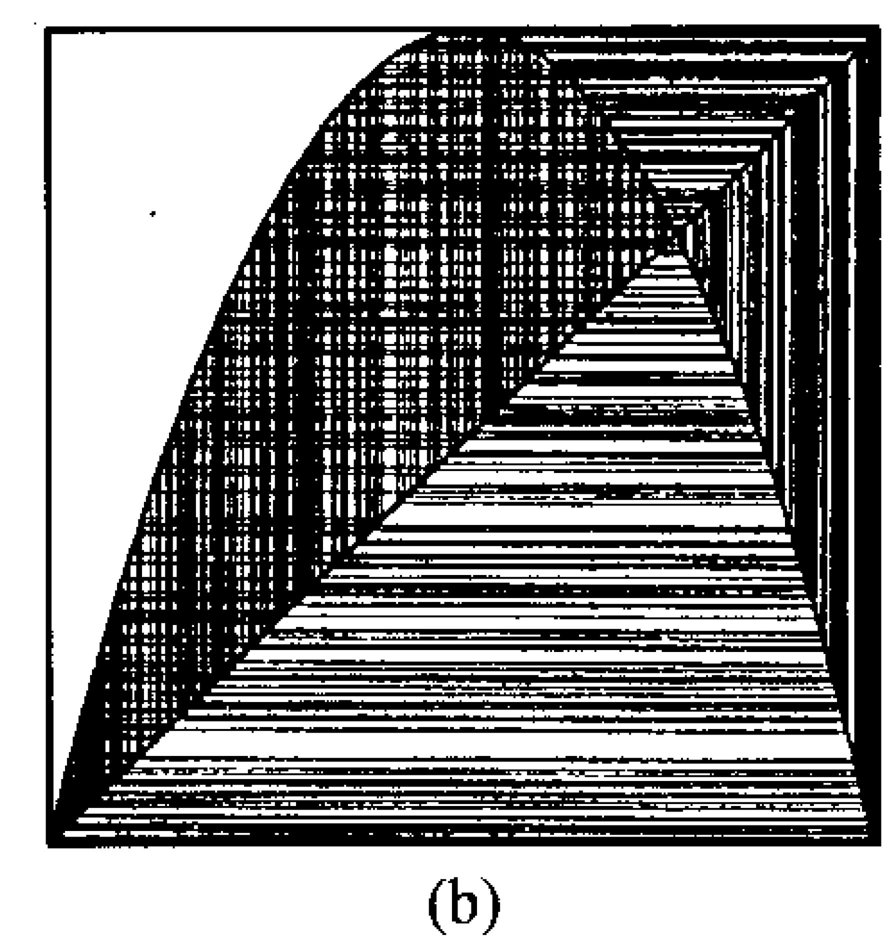

(b)

图 15.8　种子 0.123 在 f_4 下的轨道, (a) 中用了 200 次迭代, (b) 中用了 500 次迭代

15.4　混　　沌

在本节，我们介绍几个典型的混沌的一维离散动力系统例子. 回忆一下，子集 $U \subset W$ 称为在 W 中**稠密的**，如果 U 有点可以任意的接近大集合 W 中的任意点. 设 f 是一个将区间 $I = [\alpha, \beta]$ 映到自身的映射，与洛伦茨模型中一样，我们称 f 是**混沌的**，如果

(1) f 的周期点在 I 中稠密;

(2) f 在 I 上是传递的，即，对于 I 的任意两个子区间 U_1, U_2，都存在 $x_0 \in U_1$ 和一个 $n > 0$，使得 $f^n(x_0) \in U_2$;

(3) f 在 I 中有敏感依赖性，即，存在**敏感常数** $\beta > 0$，使得对于任意的 $x_0 \in I$ 以及关于 x_0 的任意开区间 U，都存在某一种子 $y_0 \in U$ 和 $n > 0$ 使得

$$|f^n(x_0) - f^n(y_0)| > \beta.$$

可以证明，传递性条件等价于存在一条在 I 中稠密的轨道. 显然，存在稠密轨道蕴涵传递性，这是因为稠密轨道会重复地经过 I 中的任意开子区间. 另一个方向的证明依赖于分析中的 Baire 纲定理，我们不打算在这里证明.

奇妙的是，对于区间上的映射，混沌定义中的条件 3 是多余的 [8]. 这多少有些令人惊奇，因为定义中的前两个条件本质上是拓扑的，而第三个则是一个度量性质 (它依赖于距离的概念).

现在我们来讨论混沌一维映射中的几个经典例子.

例　(**加倍映射**) 不连续函数 $D:[0,1)\to[0,1)$ 的定义由 $D(x)=2x \bmod 1$ 给出, 即

$$D(x)=\begin{cases}2x & \text{如果 } 0\leqslant x<1/2,\\ 2x-1 & \text{如果 } 1/2\leqslant x<1.\end{cases}$$

简单的计算表明 $D^n(x)=2^n x \bmod 1$, 从而 D^n 的图像由 2^n 条斜率为 2^n 的线段组成, 并且每条线段都延伸到整个区间 $[0,1)$(见图 15.9).

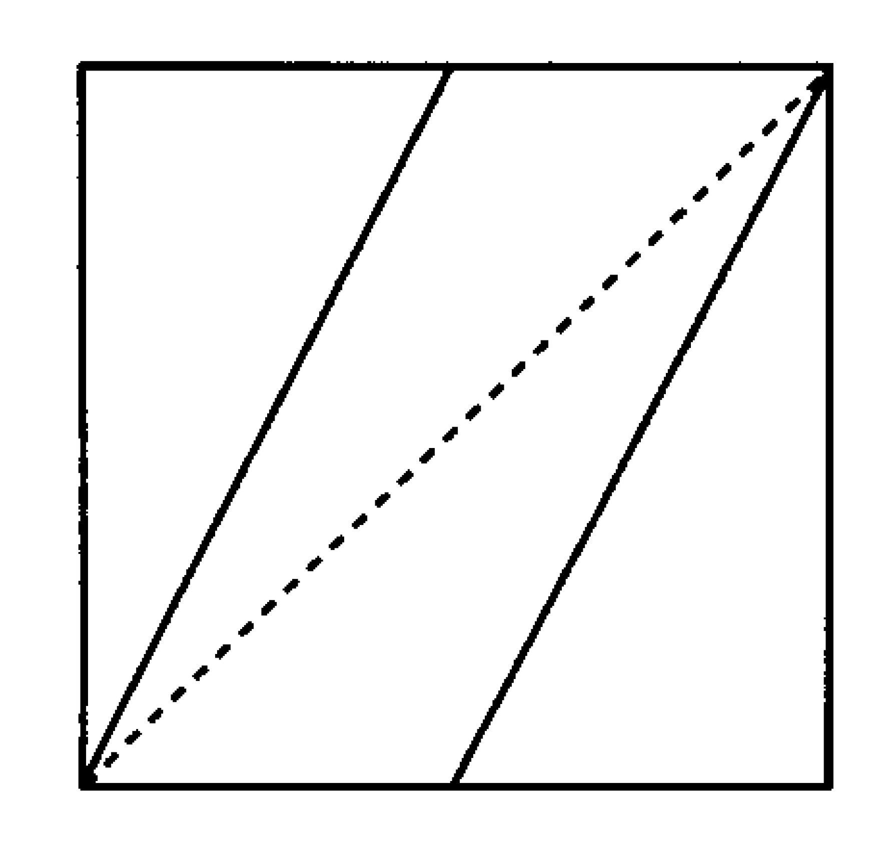
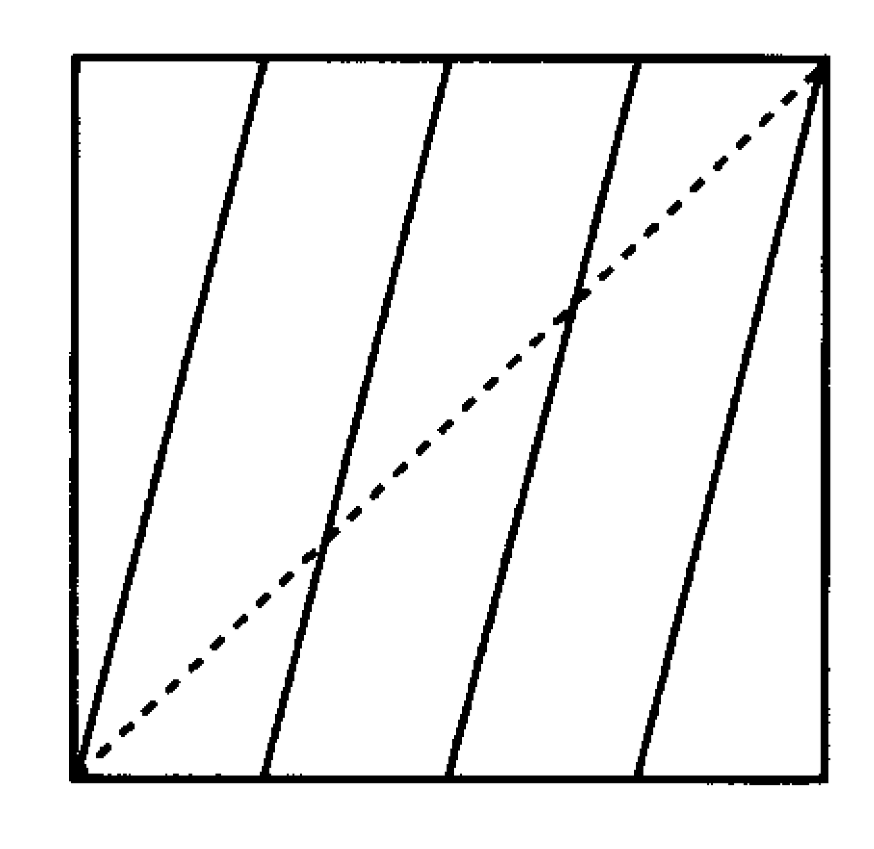
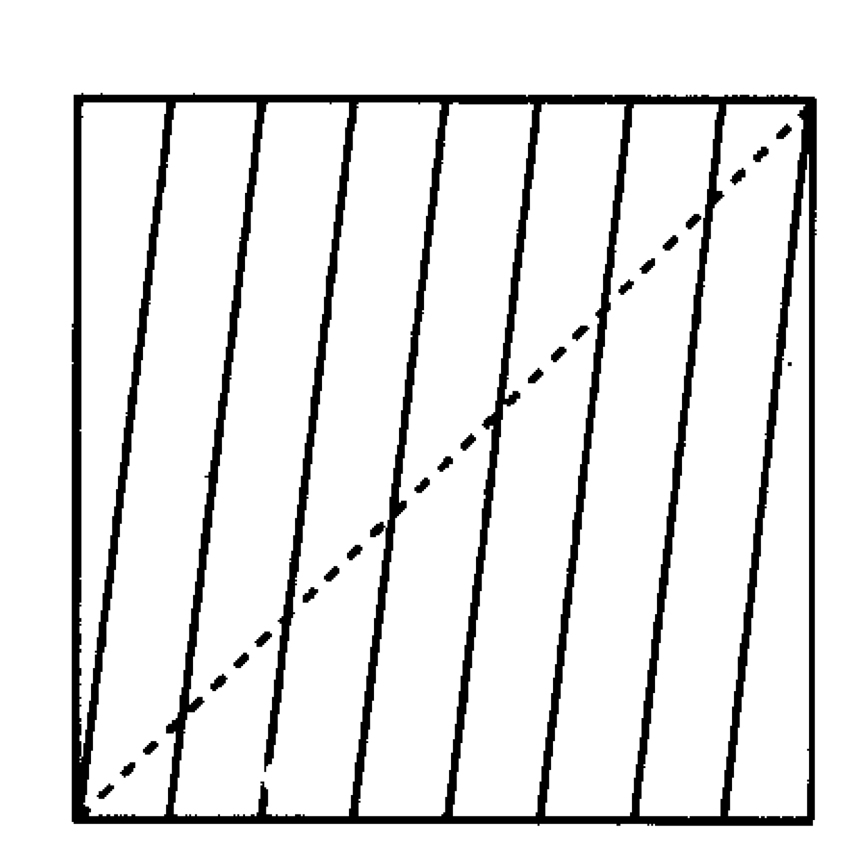




图 15.9　$[0,1)$ 上的加倍映射 D 及其高次迭代 D^2 和 D^3 的图像

为了看出加倍映射 D 在 $[0,1)$ 上是混沌的, 注意 D^n 将任何形如 $[k/2^n,(k+1)/2^n), k=0,1,\cdots,2^n-1$ 的区间映满区间 $[0,1)$. 于是 D^n 的图像与对角线 $y=x$ 在这些区间上的某点相交, 从而在每个这种区间中都有一个周期点. 由于这些区间的长度为 $1/2^n$, 因而周期点在 $[0,1)$ 中稠密. 传递性也可由此而得, 因为任给开区间 J, 对于充分大的 n, 总可以在 J 内部找到一个形如 $[k/2^n,(k+1)/2^n)$ 的区间. 从而, D^n 就将 J 映满整个 $[0,1)$. 这同时也证明了敏感性, 这里我们取敏感常数为 $1/2$ 即可.

注意, 此时可以将 D 的所有周期点明确地写出来 (见习题 5a). 有趣的是, 如果使用计算机迭代加倍映射, 将会发现所有的轨道最终都固定在 0 处, 这当然是错误的! 对于这种现象的一个解释, 见习题 5c. ■

例　(**帐篷映射**) 现在考虑和加倍映射密切相关的一个连续映射, 其定义如下:

$$T(x)=\begin{cases}2x & \text{如果 } 0\leqslant x<1/2,\\ -2x+2 & \text{如果 } 1/2\leqslant x\leqslant 1.\end{cases}$$

T 称为**帐篷映射**. 见图 15.10. 利用 T^n 的图像, T 在 $[0,1]$ 上为混沌的证明可以像加倍映射情形一样地给出 (见习题 15).

观察帐篷映射 T 的图像与 15.3 节中讨论的合理映射 $f_4(x)=4x(1-x)$ 的图像, 似乎二者在迭代下应该有许多相同的性质. 事实上, 的确如此. 为了理解这一点, 我们需要再次引入共轭的概念, 不过这次是对离散系统引入.

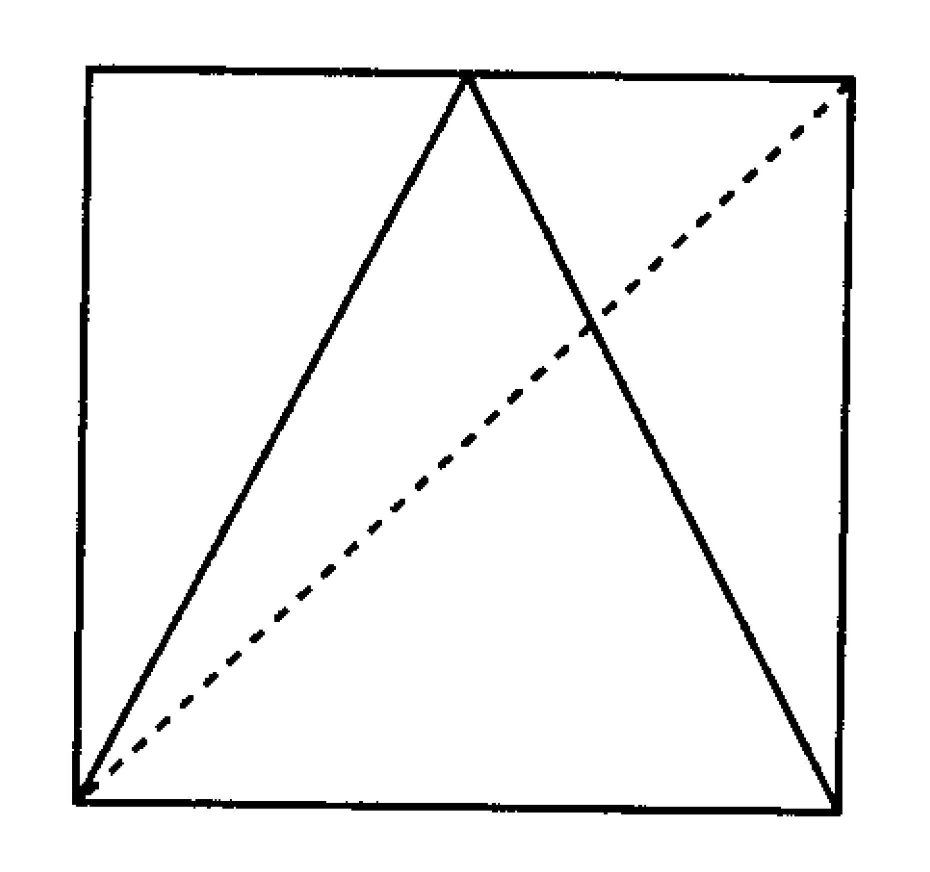
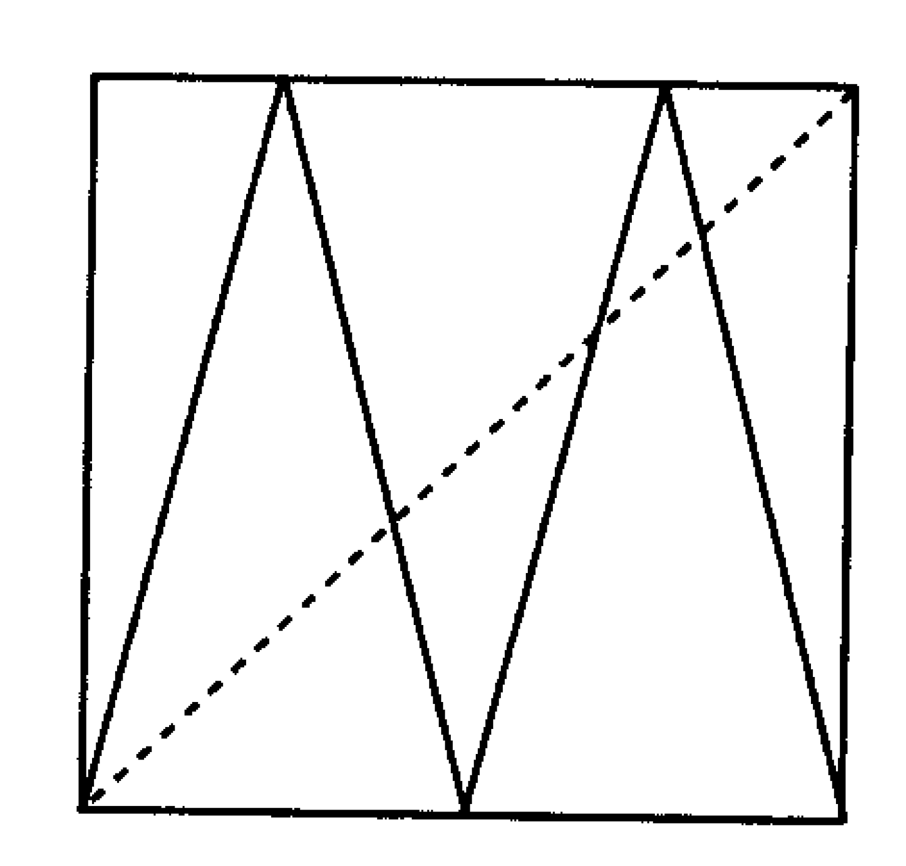
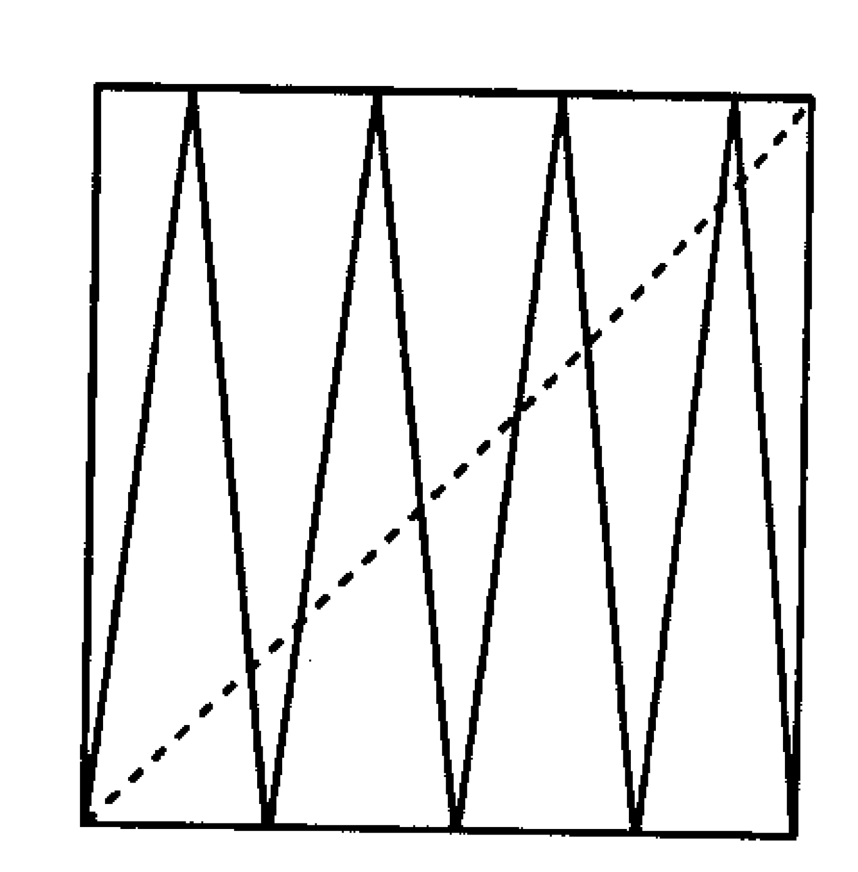




图 15.10　$[0,1]$ 上的帐篷映射 T 及其高次迭代 T^2 和 T^3 的图像

假设 I, J 是两个区间, $f: I \to I$, $g: J \to J$ 为两个映射. 我们称 f 和 g 是**共轭的**, 如果存在一个同胚 $h: I \to J$ 使得 h 满足**共轭方程** $h \circ f = g \circ h$. 同流的情形一样, 一个共轭将 f 的轨道带到 g 的轨道. 这是因为, 对所有的 $x \in I$, 我们有 $h(f^n(x)) = g^n(h(x))$, 因而 h 将 f 迭代下 x 轨道上的第 n 个点带到 g 迭代下 $h(x)$ 轨道上的第 n 个点. 同样, h^{-1} 将 g 的轨道带到 f 的轨道. ■

例　考虑合理函数 $f_4: [0,1] \to [0,1]$ 以及二次函数 $g: [-2,2] \to [-2,2]$, 其中 $g(x) = x^2 - 2$. 令 $h(x) = -4x + 2$, 可见 h 将 $[0,1]$ 带到 $[-2,2]$. 而且, 我们有 $h(4x(1-x)) = (h(x))^2 - 2$, 因而 h 满足共轭方程, 于是 f_4 和 g 是共轭的. ■

由于共轭将一个混沌系统映到另一个混沌系统, 因而从研究混沌的角度看, 共轭关系是很重要的.

命题　假设 $f: I \to I$ 和 $g: J \to J$ 通过 h 共轭, 这里 I 和 J 都是 $\mathbb{R}$ 中具有有限长的闭区间. 如果 f 在 I 上是混沌的, 则 g 在 J 上是混沌的.

证明　令 U 为 J 的一个开子区间, 考虑 $h^{-1}(U) \subset I$. 由于 f 的周期点在 I 中稠密, 因而存在 f 的周期点 $x \in h^{-1}(U)$. 设 x 的周期为 n. 则由共轭方程可得

$$g^n(h(x)) = h(f^n(x)) = h(x).$$

这就给出了 g 在 U 中的一个周期点 $h(x)$, 从而证明了 g 的周期点在 J 中稠密.

如果 U 和 V 是 J 的开子区间, 则 $h^{-1}(U)$ 和 $h^{-1}(V)$ 就是 I 的开子区间. 由 f 的传递性可知, 存在 $x_1 \in h^{-1}(U)$, 使得对于某个 m 有 $f^m(x_1) \in h^{-1}(V)$. 于是 $h(x_1) \in U$, 并且我们有 $g^m(h(x_1)) = h(f^m(x_1)) \in V$, 从而 g 也是传递的.

对于敏感性, 假设 f 具有敏感常数 β. 令 $I = [\alpha_0, \alpha_1]$, 我们不妨假设 $\beta < \alpha_1 - \alpha_0$. 对任何的 $x \in [\alpha_0, \alpha_1 - \beta]$, 考虑函数 $|h(x+\beta) - h(x)|$. 这是 $[\alpha_0, \alpha_1 - \beta]$ 上的一个连续函数, 而且是正的. 从而它有一个最小值 β'. 由此可知, h 将 I 中一个长度为 β 的区间带到 J 中一个长度至少为 β' 的区间. 容易验证 β' 就是 g 的一个敏感常数. 证毕. ■

在两个具有等价的动力学的函数之间, 并不是总可以找到共轭. 但是, 我们可以将共轭必须是一对一的这个条件放松, 而仍然保证上一个命题的结论成立. 如

果 h 是一个至多 n 到 1 的连续函数, 并且满足共轭方程 $h\circ f = g\circ h$, 则称 h 是一个 f 和 g 之间的**半共轭**. 容易验证, 半共轭也保持有限长区间上的混沌行为 (见习题 12). 半共轭不一定保持环的最小周期, 但它的确将环映到环.

例 帐篷映射 T 和合理函数 f_4 是半共轭的. 为此, 我们令

$$h(x) = \frac{1}{2}(1-\cos 2\pi x).$$

则 h 将 $[0,1]$ 区间以二对一的方式映到自身, 但是要除去 1/2 这一点, 它是唯一被映到 1 的点. 我们来计算

$$\begin{aligned} h(T(x)) &= \frac{1}{2}(1-\cos 4\pi x) \\ &= \frac{1}{2}-\frac{1}{2}(2\cos^2 2\pi x - 1) \\ &= 1-\cos^2 2\pi x \\ &= 4\left(\frac{1}{2}-\frac{1}{2}\cos 2\pi x\right)\left(\frac{1}{2}+\frac{1}{2}\cos 2\pi x\right) \\ &= f_4(h(x)). \end{aligned}$$

因而, h 是 T 和 f_4 之间的一个半共轭. 注意, 我们总可以找到任意小的子区间使得它在 T 的迭代下映满整个 $[0,1]$. 因而 f_4 将这些子区间在 h 下的像也映满整个 $[0,1]$. 由于 h 是连续的, 这些子区间的像可以取得任意小. 因而我们也可以取 1/2 作为 f_4 的一个敏感常数. 于是我们就证明了如下命题: ■

命题 合理函数 $f_4(x) = 4x(1-x)$ 在单位区间上是混沌的. ■

15.5 符号动力学

现在我们转到分析混沌系统最有用的工具之一：**符号动力学**. 这里我们仅仅给出一个如何运用符号动力学的例子, 更多的一些例子将会出现在下一章.

考虑合理映射 $f_\lambda(x) = \lambda x(1-x)$, 其中 $\lambda > 4$. 图像迭代似乎说明几乎所有的轨道都要趋于 $-\infty$(见图 15.11). 当然, 并不真的如此, 这是因为这个函数具有不动点以及其他一些周期点. 我们将看到, 事实上, 在一个意想不到的 "大" 集合 (称为**康托集**) 上都充满了这个函数的混沌行为.

与 $\lambda \leqslant 4$ 的情形不同, 当 $\lambda > 4$ 时, 区间 $I = [0,1]$ 不再是不变的：某些轨道将逃逸 I, 然后趋于 $-\infty$. 我们的目的是理解非逃逸轨道的行为. 记 Λ 是由那些轨道永不离开 I 的点所构成的集合. 如图 15.12a 所示, 存在一个开区间 A_0 使得 f_λ 在其上的取值严格大于 1. 于是, 对所有的 $x\in A_0$, 都有 $f_\lambda^2(x) < 0$, 因而 A_0 中所

有点的轨道都趋于 $-\infty$. 注意, 任何离开 I 的轨道在趋于 $-\infty$ 之前必须首先进入 A_0. 而且 A_0 的两个端点的轨道最终都固定在 0 点, 因而这两个端点包含在 Λ 中. 记 A_1 为 A_0 在 I 中的原像: A_1 由 I 中的两个开区间构成, 在 A_0 的左右两边各有一个. A_1 中的所有点都被 f_λ 映入 A_0, 因而它们的轨道也趋于 $-\infty$. 而且, A_1 的端点最终都是不动点. 如此继续下去, 我们看到, A_1 的两个开区间中每一个都包含由一对不相交的开区间构成的 A_1 的原像, 因而首次迭代位于 A_1 中的点就由 4 个开区间构成, 它们的二次迭代在 A_0 中, 于是这些点的轨道也趋于 $-\infty$. 将这 4 个区间的并记为 A_2. 一般地, 记 A_n 是 I 中的那些 n 次迭代位于 A_0 中的点构成的集合. A_n 由 I 中 2^n 个互不相交的开区间构成. 对于 I 中的任一点, 如果它的轨道离开 I, 则这个点必须位于某一个 A_n 中. 于是我们就有

$$\Lambda = I - \bigcup_{n=0}^{\infty} A_n.$$

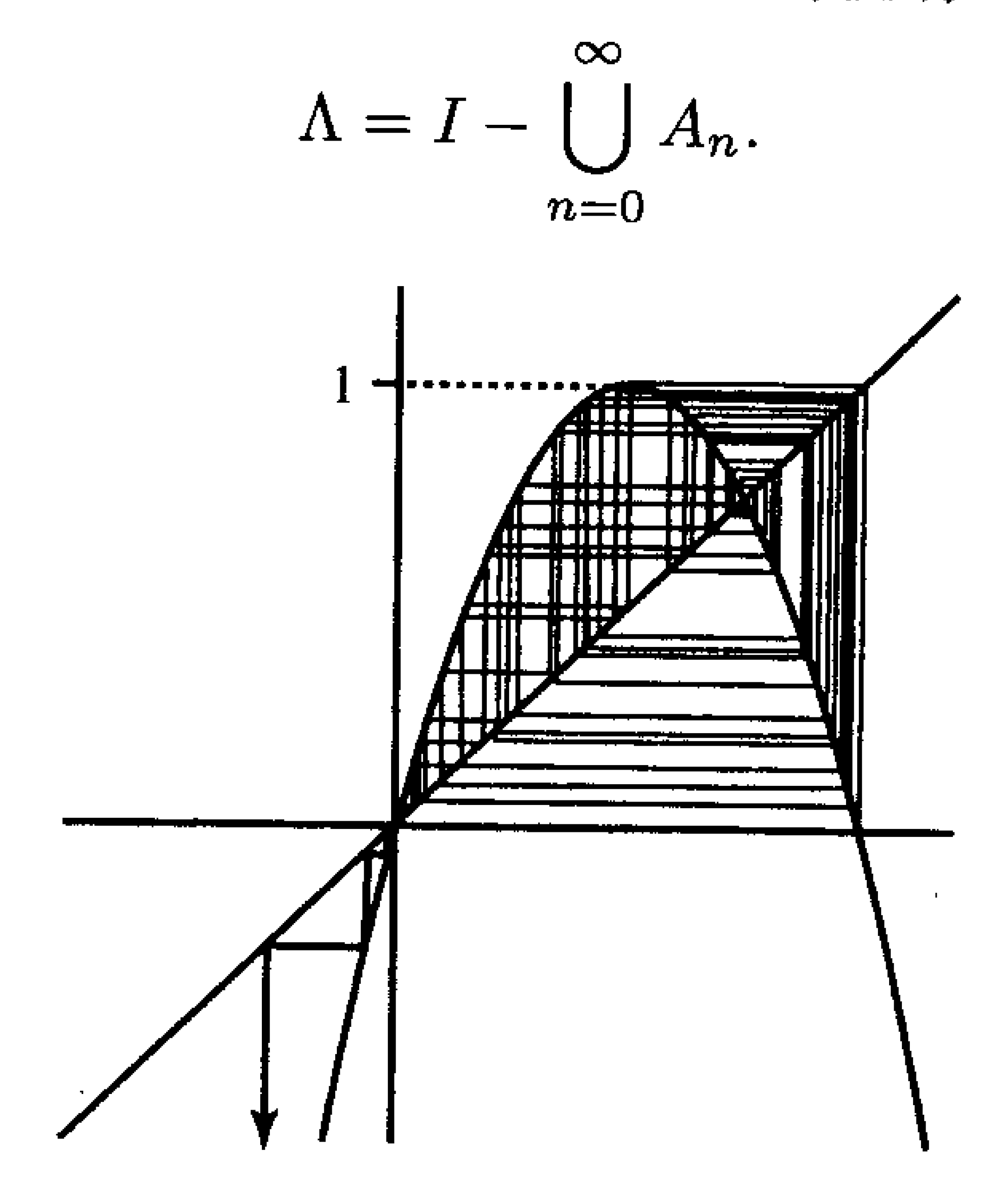


图 15.11 对于 $\lambda > 4$ 的合理函数 f_λ, 其典型轨道在围绕单位区间游荡一段时间后似乎都要趋于 $-\infty$

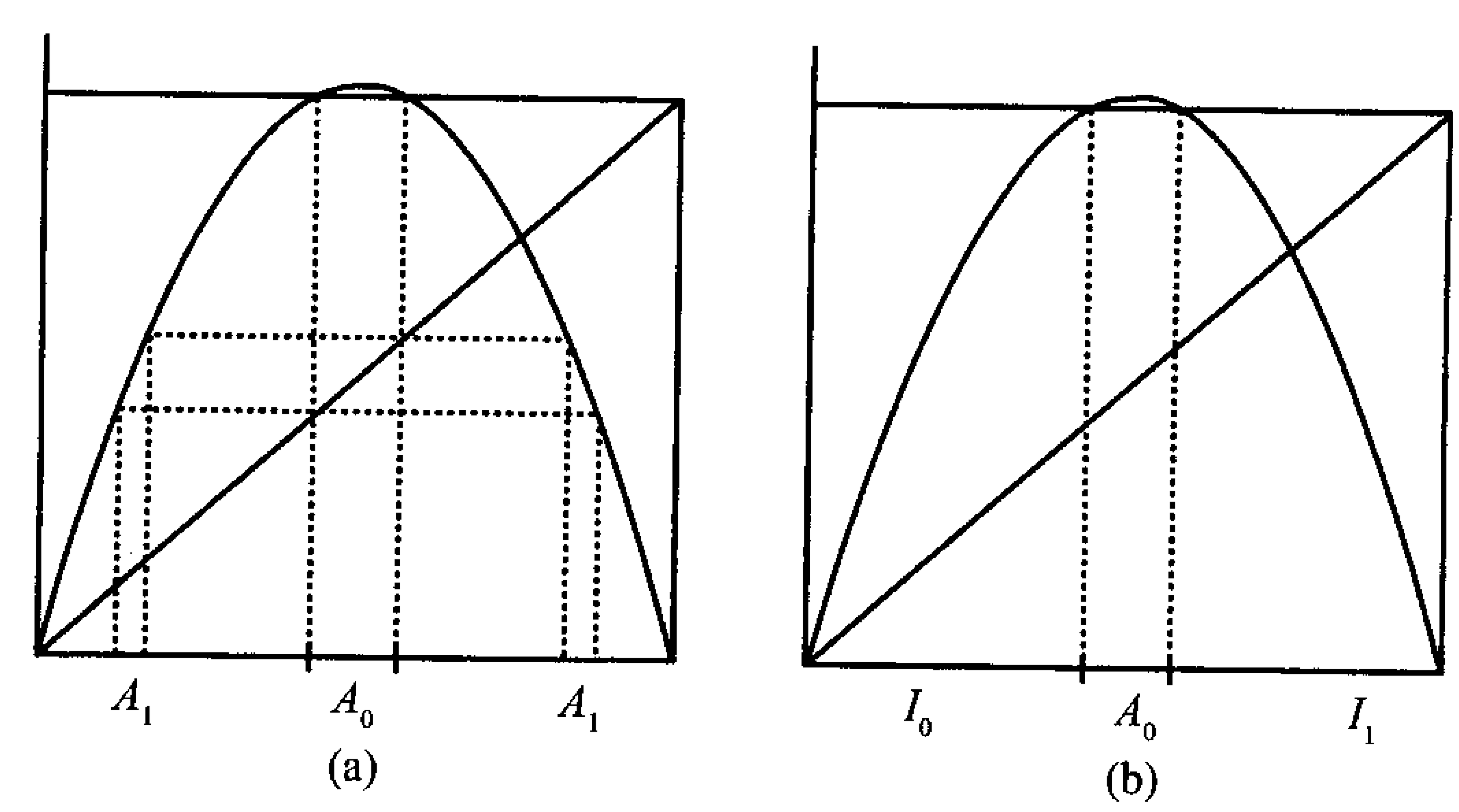


图 15.12 (a) I 中的逃逸集由一组互不相交的开区间构成. (b) 区间 I_0 和 I_1 分别位于 A_0 的左右两边

为了理解 f_λ 在 I 上的动力学, 我们引入符号动力学. 直到本节结束, 我们始终都用 I_0 和 I_1 分别记 $I - A_0$ 中的左右两个闭区间 (见图 15.12b). 给定 $x_0 \in \Lambda$, x_0 的整个轨道都位于 $I_0 \cup I_1$ 中. 于是, 我们可以将一个由 0 和 1 组成的无穷序列 $S(x_0) = (s_0 s_1 s_2 \cdots)$ 关联到点 x_0, 其关联方式如下：

$$s_j = k \quad \text{当且仅当} \quad f_\lambda^j(x_0) \in I_k.$$

即, 我们只是观察 $f_\lambda^j(x_0)$ 在 I_0 和 I_1 之间的来回跳动, 然后根据 $f_\lambda^j(x_0)$ 位于哪个区间而对第 j 个位置赋以 0 或 1. 序列 $S(x_0)$ 称为 x_0 的**旅行路线**(或**旅程**).

例　不动点 0 的旅程为 $S(0) = (000\cdots)$. I_1 中的不动点 x_λ 的旅程为 $S(x_\lambda) = (111\cdots)$. 点 $x_0 = 1$ 是最终不动的, 其旅程为 $S(1) = (1000\cdots)$. 一个在 I_0 和 I_1 之间来回跳动的 2 环的旅程为 $(\overline{01}\cdots)$ 或 $(\overline{10}\cdots)$, 其中 $\overline{01}$ 就是将块 01 重复无穷多次构成的序列.

记 Σ 为所有可能的 0, 1 序列构成的集合. 空间 Σ 中的一个“点”就是形如 $s = (s_0 s_1 s_2 \cdots)$ 的一个无穷序列. 为了便于想象空间 Σ, 我们需要知道 Σ 中两个不同的点之间的远近. 为此, 令 $s = (s_0 s_1 s_2 \cdots)$, $t = (t_0 t_1 t_2 \cdots)$ 是 Σ 中的两个点. Σ 上的一个**距离函数**, 或一个**度量**, 是满足如下 3 条的一个函数 $d = d(s,t)$:

(1) $d(s,t) \geqslant 0$, 并且 $d(s,t) = 0$ 当且仅当 $s = t$;

(2) $d(s,t) = d(t,s)$;

(3) 三角不等式：$d(s,u) \leqslant d(s,t) + d(t,u)$.

由于 Σ 并不自然地就是欧氏空间的一个子集, 我们不能在 Σ 上使用欧氏距离. 因而我们必须自己定义一个距离. 下面就是我们选择的距离函数

$$d(s,t) = \sum_{i=0}^{\infty} \frac{|s_i - t_i|}{2^i}.$$

注意这个无穷级数收敛：这个级数中的所有分子总是 0 或 1. 因而通过与几何级数比较就可知这个级数收敛：

$$d(s,t) \leqslant \sum_{i=0}^{\infty} \frac{1}{2^i} = \frac{1}{1 - 1/2} = 2.$$

可以直接验证所选的 d 满足距离函数定义中的 3 个要求 (见习题 13). 虽然这个距离函数初看起来有些复杂, 但是计算起来常常很容易. ■

例

(1) $d((\overline{0}), (\overline{1})) = \sum_{i=0}^{\infty} \frac{|0-1|}{2^i} = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2$;

(2) $d((\overline{01}), (\overline{10})) = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2$;

(3) $d\big((\overline{01}),(\overline{1})\big)=\sum\limits_{i=0}^{\infty}\dfrac{1}{4^i}=\dfrac{1}{1-1/4}=\dfrac{4}{3}$. ■

Σ 上的距离函数的重要性在于, 有了距离之后, 我们就可以知道点什么时候在接近, 什么时候在分开. 特别地, 我们有:

命题 设 $s=(s_0s_1s_2\cdots)$, $t=(t_0t_1t_2\cdots)\in\Sigma$.

(1) 如果对 $j=0,\cdots,n$ 都有 $s_j=t_j$, 则 $d(s,t)\leqslant 1/2^n$;

(2) 反之, 如果 $d(s,t)<1/2^n$, 则对 $j=0,\cdots,n$ 都有 $s_j=t_j$.

证明 在情形 (1), 我们有

$$d(s,t)=\sum_{i=0}^{n}\frac{|s_i-s_i|}{2^i}+\sum_{i=n+1}^{\infty}\frac{|s_i-t_i|}{2^i}\leqslant 0+\frac{1}{2^{n+1}}\sum_{i=0}^{\infty}\frac{1}{2^i}=\frac{1}{2^n}.$$

另一方面, 如果 $d(s,t)<1/2^n$, 则对任一 $j\leqslant n$, 必定有 $s_j=t_j$, 这是因为否则的话, 就有 $d(s,t)\geqslant |s_j-t_j|/2^j=1/2^j\geqslant 1/2^n$. ■

现在我们就在 Σ 中具有了远近的概念, 下面我们来证明本章的主要定理:

定理 当 $\lambda>4$ 时, 旅程函数 $S:\Lambda\to\Sigma$ 是一个同胚.

证明 事实上我们将只对 λ 充分大的情形证明这个定理: 我们假设存在 $K>1$, 使得对所有的 $x\in I_0\cup I_1$, 都有 $|f_\lambda'(x)|>K>1$. 读者可以验证, 只要 $\lambda>2+\sqrt{5}$ 就足够了. 当 $4<\lambda\leqslant 2+\sqrt{5}$ 时, 证明要复杂些, 参见 [25].

我们首先证明 S 是一对一的. 令 $x,y\in\Lambda$, 并且假设 $S(x)=S(y)$. 于是, 对每个 n, $f_\lambda^n(x), f_\lambda^n(y)$ 都位于 1/2 的同一边. 这意味着 f_λ 在 $f_\lambda^n(x)$ 和 $f_\lambda^n(y)$ 之间的区间上是单调的. 于是, 这个区间在 f_λ 下的像仍然在 $I_0\cup I_1$ 中. 由于在这个区间中的所有点处, 都有 $|f_\lambda'|>K>1$, 因而, 同 15.1 节中一样, f_λ 每迭代一次, 就将这个区间的长度扩张 K 倍. 于是 $f_\lambda^n(x)$ 和 $f_\lambda^n(y)$ 之间的距离将无限增长, 因而这两个点最终必定要位于 A_0 的两边. 这与它们有相同的旅程矛盾.

为了看出 S 是满的, 我们先引入下面的记号. 令 $J\subset I$ 为一个闭区间. 记

$$f_\lambda^{-n}(J)=\{x\in I|f_\lambda^n(x)\in J\},$$

即 $f_\lambda^{-n}(J)$ 是 J 在 f_λ^n 下的原像. 从 $\lambda>4$ 时 f_λ 的图像可知, 如果 $J\subset I$ 是一个闭区间, 则 $f_\lambda^{-1}(J)$ 由两个闭子区间构成, 一个在 I_0 中, 另一个在 I_1 中.

现在令 $s=(s_0s_1s_2\cdots)$. 我们需要构造一个 $x\in\Lambda$ 使得 $S(x)=s$. 为此, 我们定义

$$\begin{aligned}I_{s_0s_1\cdots s_n}&=\{x\in I|x\in I_{s_0}, f_\lambda(x)\in I_{s_1},\cdots,f_\lambda^n(x)\in I_{s_n}\}\\&=I_{s_0}\cap f_\lambda^{-1}(I_{s_1})\cap\cdots\cap f_\lambda^{-n}(I_{s_n}).\end{aligned}$$

我们断言, $I_{s_0s_1\cdots s_n}$ 构成了一个非空的闭区间套序列. 注意,

$$I_{s_0s_1\cdots s_n} = I_{s_0} \cap f_\lambda^{-1}(I_{s_1\cdots s_n}).$$

根据归纳法, 我们不妨假设 $I_{s_1\cdots s_n}$ 是一个非空的子区间, 于是, 由上面的观察, $f_\lambda^{-1}(I_{s_1\cdots s_n})$ 由两个闭区间构成, 一个在 I_0 中, 另一个在 I_1 中, 因而 $I_{s_0}\cap f_\lambda^{-1}(I_{s_1\cdots s_n})$ 就是一个闭区间. 这些区间是嵌套的, 这是因为

$$I_{s_0s_1\cdots s_n} = I_{s_0s_1\cdots s_{n-1}} \cap f_\lambda^{-n}(I_{s_n}) \subset I_{s_0s_1\cdots s_{n-1}}.$$

由此可得

$$\bigcap_{n=0}^{\infty} I_{s_0s_1\cdots s_n}$$

是非空的. 注意, 如果 $x \in \cap_{n\geqslant 0} I_{s_0s_1\cdots s_n}$, 则 $x \in I_{s_0}, f_\lambda(x) \in I_{s_1}$, 等等. 因而, $S(x) = (s_0s_1\cdots)$. 这就证明了 S 是满的.

注意, $\cap_{n\geqslant 0} I_{s_0s_1\cdots s_n}$ 只包含唯一的点. 从 S 是一对一的事实可以立刻推得此事. 特别地, 当 $n \to \infty$ 时, $I_{s_0s_1\cdots s_n}$ 的长度趋于 0.

为了证明 S 的连续性, 我们任取 $x \in \Lambda$, 假设 $S(x) = (s_0s_1s_2\cdots)$. 令 $\epsilon > 0$, 选取 n 使得 $1/2^n < \epsilon$. 对于所有可能的排列 $t_0t_1\cdots t_n$, 考虑上面定义的闭子区间 $I_{t_0t_1\cdots t_n}$. 这些子区间都是互不相交的, 而且 Λ 就包含在它们的并中. 共有 2^{n+1} 个这样的子区间, 且 $I_{s_0s_1\cdots s_n}$ 就是其中的一个. 因而可取 δ 使得当 $y \in \Lambda$ 且 $|x-y| < \delta$ 时, 就蕴涵 $y \in I_{s_0s_1\cdots s_n}$. 于是, $S(y)$ 的前 $n+1$ 项与 $S(x)$ 的相同, 从而根据上一命题, 我们有

$$d(S(x), S(y)) \leqslant \frac{1}{2^n} < \epsilon.$$

这就证明了 S 的连续性. 容易验证 S^{-1} 也是连续的, 从而 S 是一个同胚. ■

15.6 移位映射

现在构造一个映射 $\sigma: \Sigma \to \Sigma$, 它具有下面的性质:

(1) σ 是混沌的;

(2) σ 与 f_λ 在 Λ 上的限制是共轭的;

(3) 从动力系统的角度看, σ 可以被彻底地理解.

随着我们的进行, 最后一条的含义将会变得清晰.

定义**移位映射**$\sigma: \Sigma \to \Sigma$ 为

$$\sigma(s_0s_1s_2\cdots) = (s_1s_2s_3\cdots).$$

即, 移位映射的作用就是将 Σ 中每个序列的第一个数字扔掉. 注意, σ 是 Σ 上的一个二对一的满射, 这是因为, 如果 $(s_0s_1s_2\cdots)\in\Sigma$, 则

$$\sigma(0s_0s_1s_2\cdots)=\sigma(1s_0s_1s_2\cdots)=(s_0s_1s_2\cdots).$$

命题 移位映射 $\sigma:\Sigma\to\Sigma$ 是连续的.

证明 令 $s=(s_0s_1s_2\cdots)\in\Sigma$, $\epsilon>0$. 选取 n 使得 $1/2^n<\epsilon$. 令 $\delta=1/2^{n+1}$. 假设 $d(s,t)<\delta$, 其中 $t=(t_0t_1t_2\cdots)$. 则对 $i=0,1,\cdots,n+1$, 有 $s_i=t_i$.

现在 $\sigma(t)=(s_1s_2\cdots s_ns_{n+1}t_{n+2}\cdots)$, 从而 $d(\sigma(s),\sigma(t))\leqslant 1/2^n<\epsilon$. 这就证明了 σ 是连续的. ■

注意, 我们可以容易地写出移位映射的任一周期的所有周期点. 事实上, 不动点就是 $(\overline{0})$ 和 $(\overline{1})$. 2 环就是 $(\overline{01})$ 和 $(\overline{10})$. 一般地, 周期为 n 的周期点就是无穷次重复长度为 n 的块构成: $(\overline{s_0\cdots s_{n-1}})$. 注意, 相比于 f_λ, σ 是多么地好: 不信, 你哪天试试将 f_λ 的所有周期为 n 的周期点明确地写出来! 但是, 它们的确就在那里, 而且我们还知道它们的大致位置, 因为我们有:

定理 旅程函数 $S:\Lambda\to\Sigma$ 给出了 f_λ 和移位映射 σ 之间的一个共轭.

证明 在上节, 我们证明了 σ 是一个同胚. 因而只需证明 $S\circ f_\lambda=\sigma\circ S$. 为此, 令 $x_0\in\Lambda$, 并假设 $S(x_0)=(s_0s_1s_2\cdots)$. 则我们有 $x_0\in I_{s_0}, f_\lambda(x_0)\in I_{s_1}, f_\lambda^2(x_0)\in I_{s_2}$, 等等. 于是就有 $f_\lambda(x_0)\in I_{s_1}, f_\lambda^2(x_0)\in I_{s_2}$, 等等. 这也就是说 $S(f_\lambda(x_0))=(s_1s_2s_3\cdots)$, 即 $S(f_\lambda(x_0))=\sigma(S(x_0))$, 这正是我们要证的. ■

现在, 我们不仅可以写出 σ 的所有周期点, 而且还可以明确地写出 Σ 中的一个点, 它的轨道是稠密的. 下面就是这样的一个点:

$$s^*=(\underbrace{0\ 1}_{\text{长为 1 的块}}\ |\ \underbrace{00\ 01\ 10\ 11}_{\text{长为 2 的块}}\ |\ \underbrace{000\ 001\ \cdots}_{\text{长为 3 的块}}\ |\ \underbrace{\cdots}_{\text{长为 4 的块}}).$$

依次列出 0,1 构成的所有长度为 1 的块、长度为 2 的块、长度为 3 的块, 等等, 就构造出了序列 s^*. 显然, 存在 σ 在 s^* 上的某次迭代使得所得到序列与 Σ 中任一给定的序列在任意长的初始位置上都相同. 即, 任给 $t=(t_0t_1t_2\cdots)\in\Sigma$, 都可以找到 k, 使得序列 $\sigma^k(s^*)$ 起始的 $n+1$ 个数字为 $t_0t_1\cdots t_n$, 即

$$\sigma^k(s^*)=(t_0t_1\cdots t_ns_{n+1}s_{n+2}\cdots),$$

从而

$$d(\sigma^k(s^*),t)\leqslant 1/2^n.$$

因而 s^* 的轨道可以任意接近 Σ 中的每一个点. 这就证明了 s^* 在 σ 下的轨道在 Σ 中稠密, 因此 σ 是传递的. 注意, 简单地将 s^* 中的那些块重新排序, 我们就可以得

到许多其他的点, 它们的轨道也在 Σ 中稠密. 假如我们要确定一个种子, 使得它在一个像 f_4 这样的二次函数下的轨道在 $[0,1]$ 中稠密, 大家想象一下, 这将会是多么地困难! 这就是我们早先所说的 σ 的动力学可以被 "彻底地理解" 的含义.

移位映射也具有敏感依赖性. 事实上, 我们可以将敏感常数就取成 2, 即取成 Σ 中两点之间可能的最大距离. 原因如下: 如果 $s = (s_0 s_1 s_2 \cdots) \in \Sigma$, 记 $\hat{s}_j$ 为 "非 s_j" (即, 如果 $s_j = 0$, 则 $\hat{s}_j = 1$; 如果 $s_j = 1$, 则 $\hat{s}_j = 0$), 于是点 $s' = (s_0 s_1 \cdots s_n \hat{s}_{n+1} \hat{s}_{n+2} \cdots)$ 满足

(1) $d(s, s') = 1/2^n$, 但是

(2) $d(\sigma^{n+1}(s), \sigma^{n+1}(s')) = 2$.

作为一个推论, 我们就证明了:

定理　移位映射 σ 在 Σ 上是混沌的, 因而根据上个定理中的共轭关系, 合理映射 f_λ 在 $\lambda > 4$ 时是混沌的. ■

虽然 f_λ 在 Λ 上是混沌的, 然而符号动力学却为集 Λ 上 f_λ 的动力学提供了一个可计算的模型.

15.7　三分康托集

前面我们曾提到过, Λ 是一个康托集的例子. 这里, 我们将描述这种集合的一个最简单的例子, **三分康托集** C. 我们将看到, 这个集合具有一些出乎意料的有趣性质.

为了定义 C, 我们从单位闭区间 $I = [0,1]$ 开始. 规则是: 每当看到一个闭区间, 我们就将中间三分之一的开区间拿掉. 于是, 在第一步, 我们拿掉 $(1/3, 2/3)$, 留下 2 个闭区间 $[0, 1/3]$ 和 $[2/3, 1]$. 然后重复这一步, 再拿掉留下的 2 个闭区间的中间三分之一. 这样我们就留下 4 个闭区间 $[0,1/9], [2/9,1/3], [2/3,7/9], [8/9,1]$. 再拿掉这些区间的中间三分之一, 就留下 2^3 个闭区间, 每个的长度都是 $1/3^3$. 以这种方式继续下去, 在第 n 步, 我们留下 2^n 个闭区间, 每个的长度都是 $1/3^n$. 当我们将这个过程进行到 $n \to \infty$ 时的极限, 最后留下的就是三分康托集 C. 注意, 这个构造与 15.5 节中 Λ 的构造非常相似! 事实上, 可以证明, Λ 同胚于 C(见习题 16 和习题 17).

在拿掉所有这些开区间后, I 中哪些点会留在 C 中呢? 当然, 0 和 1 都留在 C 中, 还有第一次拿掉的区间的端点 1/3 和 2/3 也留在 C 中. 事实上, 每个被拿掉的开区间的端点都位于 C 中, 这是因为这些点始终都不会在中间三分之一的开子区间中. 乍看起来, 似乎这些就是康托集 C 中仅有的点, 但事实远非如此. 事实上, C 中的大多数点都**不是**端点!

为了看出这一点, 我们将 C 中的每个点贴上一个地址. 这些地址是由 L, R 组

成的无穷长的串, 其定义方式如下. 在构造中的每一步, C 中的点都位于两个小的闭子区间的某一个中, 这两个小的闭子区间一个在被拿掉的开区间的左边, 一个在右边. 从而在第 n 步, 根据这个点是位于这一步被拿掉区间的左边或右边, 我们将分别赋以 L 或 R. 例如, 我们将 $LLL\cdots$ 赋给 0, 而将 $RRR\cdots$ 赋给 1. 端点 1/3 和 2/3 的地址分别是 $LRRR\cdots$ 和 $RLLL\cdots$. 在下一步, 1/9 的地址是 $LLRRR\cdots$, 这是因为在开始两步 1/9 位于 $[0,1/3]$ 和 $[0,1/9]$ 中, 此后它总是在右边的区间. 同样, 2/9 的地址是 $LRLLL\cdots$, 而 7/9 和 8/9 地地址分别是 $RLRRR\cdots$ 和 $RRLLL\cdots$.

请注意 C 的每个端点的情况. 上面的几个例子表明, 一个端点的地址总是以全部由 L 或 R 构成的无穷串结束. 但是, 对于 C 中的点, 还有很多其他可能的地址. 例如, C 中存在一个点, 它的地址是 $LRLRLR\cdots$. 这个点位于如下的交集中:

$$[0,1/3]\cap[2/9,1/3]\cap[2/9,7/27]\cap[20/81,7/27]\cdots.$$

注意, 这个点位于一个长度依次为 $1/3^n$ 的闭区间套的交集中, 而且它是这个交集中的唯一的点. 因为典型的地址都不是以全由 L, 或全由 R 构成的串结束, 这就说明 C 中的大多数点都不是端点.

事实上我们可以说得更多些: 三分康托集包含不可数多个点. 回忆一下, 一个无穷集合是**可数的**, 如果它可以和自然数之间建立一一对应; 否则这个集合就是**不可数的**.

命题 三分康托集是不可数的.

证明 假设 C 是可数的. 这意味着我们可用某种方式将 C 中的每个点都对应到一个自然数, 比如,

$$\begin{array}{ccl} 1 & : & LLLLL\cdots \\ 2 & : & RRRR\cdots \\ 3 & : & LRLR\cdots \\ 4 & : & RLRL\cdots \\ 5 & : & LRRLRR\cdots \end{array}$$

等等. 但是, 现在我们考虑这样的一个地址, 它的第一个元素是序列 1 的第一个元素的对面, 它的第二个元素是序列 2 的第二个元素的对面, 如此下去. 这是一个新的 L,R 序列 (比如, 在上面列出的例子中, 它的开始几个字母是 $RLRRL\cdots$). 于是, 我们就构造出了一个 L,R 的序列, 它与我们列出的第 n 个序列在第 n 个位置不同. 于是, 这个序列就没有被列出, 因而这个与自然数的一一对应是失败的. 这一矛盾就证明了结果. ■

我们可以用一种更熟悉的方式确定三分康托集中的点. 为此, 我们将 C 中点的地址从 L,R 的序列变换成 0, 2 的序列, 从而得到它的一个新地址; 即, 将每个 L

替换成 0, 而将每个 R 替换成 2. 为了确定 C 中一个点 x 的数值, 我们从 x 的下方来逼近它: 开始时为 0, 然后对每个 $n = 1, 2, 3, \cdots$, 向右移动 $s_n/3^n$ 个单位, 其中 $s_n = 0$ 或 2 是 x 的新地址中的第 n 个数字.

例如, 1 的老地址是 $RRR\cdots$, 新地址是 $222\cdots$, 于是 1 可以写成

$$\frac{2}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = \frac{2}{3}\sum_{n=0}^{\infty}\frac{1}{3^n} = \frac{2}{3}\left(\frac{1}{1-1/3}\right) = 1.$$

类似地, 1/3 的老地址是 $LRRR\cdots$, 新地址是 $0222\cdots$, 因而

$$\frac{0}{3} + \frac{2}{3^2} + \frac{2}{3^3} + \cdots = \frac{2}{9}\sum_{n=0}^{\infty}\frac{1}{3^n} = \frac{2}{9}\left(\frac{1}{1-1/3}\right) = \frac{1}{3}.$$

最后, 老地址是 $LRLRLR\cdots$, 新地址是 $020202\cdots$ 的点是

$$\frac{0}{3} + \frac{2}{3^2} + \frac{0}{3^3} + \frac{2}{3^4} + \cdots = \frac{2}{9}\sum_{n=0}^{\infty}\frac{1}{9^n} = \frac{2}{9}\left(\frac{1}{1-1/9}\right) = \frac{1}{4}.$$

注意, 这个点就是早先提到的 C 中的一个非端点.

敏锐的读者可能已经意识到 C 的一个点 x 的新地址就是 x 的**三进制表示**. 一个点 $x \in I$ 的三进制表示为 $a_1a_2a_3\cdots$, 如果

$$x = \sum_{i=1}^{\infty}\frac{a_i}{3^i},$$

其中 a_i 取 0、1 或 2. 于是我们看到, 在三分康托集中的点的三进制表示中, 数字 1 始终不出现.

这里我们必须要小心一些. 1/3 的三进制表示是 $1000\cdots$. 但上面我们已经看到, 1/3 的三进制表示还可以写成 $0222\cdots$. 从而 1/3 有一个不包含数字 1 的三进制表示. 事实上, C 中每个端点都有类似的两个三进制表示, 其中一个不包含数字 1.

我们已经证明 C 包含不可数多个点, 但是我们可以说得更多:

命题　三分康托集与区间 $[0,1]$ 包含同样多的点.

证明　C 由所有这样的一些点构成, 它们的三进制表示 $a_1a_2a_3\cdots$ 中只包含 0 或 2 两个数字. 取一个这样的表示, 并将其中的每个数字 2 用 1 替换, 然后将这个串看成一个二进制表示. 这样, 我们就得到了所有可能的二进制表示. 由于 $[0,1]$ 中的点都有一个二进制表示, 于是我们就得到了 C 中的点和 $[0,1]$ 中的点之间的一个对应 (至多是 2 对 1 的). ■

最后, 我们还注意到:

命题　三分康托集的长度为 0.

证明 我们先将每一步中拿掉的区间的长度相加, 这样就得到了 C 的余集的长度, 然后由此得出 C 的 "长度". 这些被拿掉的区间的长度依次为 $1/3, 2/9, 4/27, \cdots$, 因而 $I-C$ 的长度就是

$$\frac{1}{3}+\frac{2}{9}+\frac{4}{27}+\cdots=\frac{1}{3}\sum_{n=0}^{\infty}\left(\frac{2}{3}\right)^n=1.$$ ■

由于 C 只是由一些 "稀疏的" 点组成, 上面的事实并不令人惊讶. 但是, 现在我们来考虑五分康托集, 其构造与 C 类似, 只不过现在拿掉的是每个闭区间中的五分之一的开区间. 这个集合的长度非零, 但它仍然与 C 同胚. 正如我们前面所说, 这些康托集具有出乎意料的有趣性质! 而且, 请注意 f_4 在其上为混沌的集合 Λ 正是这样一种对象.

15.8 探索：立方混沌

在这个探索中, 你将研究由一族三次函数 $f_\lambda(x)=\lambda x-x^3$ 给出的离散动力系统的行为. 你必须尽量严格证明下面列出的每件事.

(1) 对所有的 $\lambda<-1$, 描述这族函数的动力学.

(2) 描述在 $\lambda=-1$ 处出现的分岔. 提示：注意, f_λ 是一个奇函数. 特别地, 当 f_λ 的图像穿过直线 $y=-x$ 时, 发生了什么?

(3) 描述在 $-1<\lambda<1$ 时 f_λ 的动力学.

(4) 描述在 $\lambda=1$ 时出现的分岔.

(5) 找出 λ 的一个取值 λ^* 使得 f_{λ^*} 有一对不变区间 $[0,\pm x^*]$, 并且在每个这种区间上, f_λ 的行为都类似于合理函数 $f_4=4x(1-x)$.

(6) 当 λ 增加地穿过 λ^* 时, 描述动力学上出现的变化.

(7) 描述当 λ 非常大时 f_λ 的动力学. 在这种情形, 描述轨道不逃逸到 $\pm\infty$ 的点构成的集合 Λ_λ.

(8) 当 λ 非常大时, 运用符号动力学建立一个序列空间以及一个相应的移位映射, 并证明 f_λ 在 Λ_λ 上是混沌的.

(9) 找到一个参数值 $\lambda'>\lambda^*$, 使得上面的两个观察结果成立.

(10) 描述当 λ 增加地穿过 λ' 时出现的分岔.

15.9 探索：轨道图

不同于上个探索, 这个探索主要是试验性的. 设计这个探索是为了让你熟悉, 当参数从 0 增加到 4 时, 合理族中丰富的动力学. 利用一台计算机以及任何看似合

适的软件, 按下面的方式构造合理族 $f_\lambda(x)=\lambda x(1-x)$ 的**轨道图**: 选取一个 N, 并且 N 等分参数空间 $[0,4]$, 得到 N 个 λ 的值, $\lambda_1,\lambda_2,\cdots,\lambda_N$. 例如, 取 $N=800$, 令 $\lambda_j=0.005j$. 对每个 λ_j 计算 0.5 在 f_{λ_j} 下的轨道, 并按照下面的方式画出这些轨道.

取水平轴作为 λ 轴, 竖直轴作为 x 轴. 在每个 λ_j 上方, 描出点 $(\lambda_j, f_{\lambda_j}^k(0.5))$, 其中 (比如说)$50\leqslant k\leqslant 250$. 即, 算出 0.5 在 f_{λ_j} 下的轨道的前 250 个点, 但是在竖直线 $\lambda=\lambda_j$ 上只显示后面的 200 个点. 实际上, 你这样显示的就是 0.5 的轨道的"命运".

你需要将这个图中的某些部分放大; 图 15.13 所显示的就是这样的一种放大图, 其中我们只显示了 λ 位于区间 $[3,4]$ 时的轨道图.

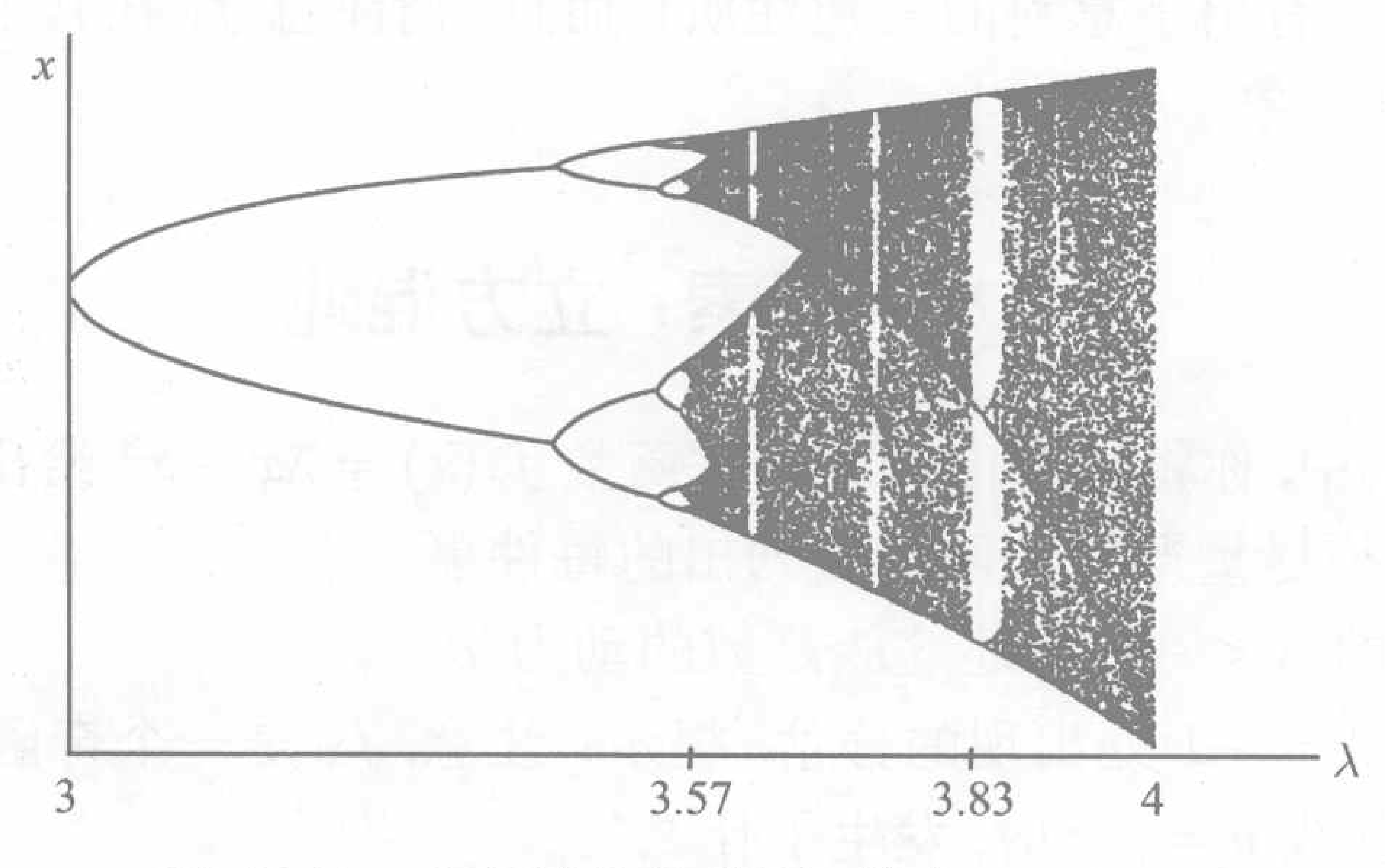


图 15.13　合理族的轨道图, 其中 $3\leqslant\lambda\leqslant 4$

(1) 由 $0\leqslant\lambda<3.57\cdots$ 所界定的区域称为**周期 1 窗口**. 描述当 λ 在这个窗口中增加时你所看到的情况. 出现了什么类型的分岔?

(2) 在上一问中的分岔附近, 你有时会看到一些模糊的点, 这是什么引起的?

(3) 观察大致由 $3.828\cdots<\lambda<3.857\cdots$ 所界定的**周期 3 窗口**. 研究当 λ 增加时产生这个窗口的分岔.

(4) 还有很多其他的周期 n 窗口 (如此称呼是由于 n 是在这个窗口中观察到的环的最小周期). 当 λ 增加时, 讨论你能找到的窗口排列方式. 特别地, 如果你放大周期 1 窗口和周期 3 窗口之间的部分, 在依次放大的过程中, 较大的窗口是如何排列的?

(5) 在这个轨道图中, 你可以观察到一些"较黑的"曲线. 它们是什么? 为什么会出现这样的情况?

习　题

1. 找出下面每个映射的所有周期点, 并指出它们是吸引的还是排斥的, 或者都不

是.

(a) $Q(x) = x - x^2$,　　(b) $Q(x) = 2(x - x^2)$,

(c) $C(x) = x^3 - \frac{1}{9}x$,　　(d) $C(x) = x^3 - x$,

(e) $S(x) = \frac{1}{2}\sin x$,　　(f) $S(x) = \sin x$,

(g) $E(x) = \mathrm{e}^{x-1}$,　　(h) $E(x) = \mathrm{e}^x$,

(i) $A(x) = \arctan x$,　　(j) $A(x) = -\frac{\pi}{4}\arctan x$.

2. 讨论下面的映射族在指定的参数值处的分岔.
   (a) $S_\lambda(x) = \lambda \sin x$,　$\lambda = 1$;
   (b) $C_\mu(x) = x^3 + \mu x$,　$\mu = -1$; (提示：利用 C_μ 是一个奇函数.)
   (c) $G_\nu(x) = x + \sin x + \nu$,　$\nu = 1$;
   (d) $E_\lambda(x) = \lambda \mathrm{e}^x$,　$\lambda = 1/\mathrm{e}$;
   (e) $E_\lambda(x) = \lambda \mathrm{e}^x$,　$\lambda = -\mathrm{e}$;
   (f) $A_\lambda(x) = \lambda \arctan x$,　$\lambda = 1$;
   (g) $A_\lambda(x) = \lambda \arctan x$,　$\lambda = -1$.
3. 考虑线性映射 $f_k(x) = kx$. 证明在参数空间中存在 4 个开集, 使得在每个开集上, f_k 的轨道的行为是类似的. 描述在那些例外情形出现的情况.
4. 对于定义在 $\mathbb{R}$ 上的函数 $f_\lambda(x) = \lambda x(1-x)$,
   (a) 描述当 $\lambda = -1$ 和 $\lambda = 3$ 时出现的分岔;
   (b) 找出所有周期为 2 的周期点;
   (c) 描述当 $\lambda = -1.75$ 时出现的分岔.
5. 对于 $[0,1)$ 上的加倍函数 D,
   (a) 明确地列出所有的周期点;
   (b) 列出轨道终止于 0 的所有点 (因而这些点是最终不动的);
   (c) 令 $x \in [0,1)$, 并假设它的二进制表示为 $a_0a_1a_2\cdots$, 其中每个 a_j 都是 0 或 1. 首先给出 $D(x)$ 的一个二进制表示. 然后解释, 为什么用计算机生成 D 的轨道时总是最终固定在 0.
6. 证明：如果 x_0 位于一个周期为 n 的环上, 则

$$(f^n)'(x_0) = \prod_{i=0}^{n-1} f'(x_i).$$

并由此得出, 对于所有的 $j = 1, \cdots, n-1$, 都有

$$(f^n)'(x_0) = (f^n)'(x_j).$$

7. 证明：如果 f_{λ_0} 在 x_0 处有一个不动点，且 $|f'_{\lambda_0}(x_0)| > 1$，则存在关于 x_0 的区间 I，以及关于 λ_0 的区间 J，使得当 $\lambda \in J$ 时，f_λ 在 I 中就有唯一不动的源点，并且它没有其他的轨道整个地位于 I 中.
8. 通过如下的两步，验证函数族 $f_c(x) = x^2 + c$ 在 $c = -3/4$ 处经历了一个倍周期分岔：
   (a) 明确地算出周期为 2 的轨道;
   (b) 证明当 $-5/4 < c < -3/4$ 时这个轨道是吸引的.
9. 仍然考虑函数族 $f_c(x) = x^2 + c$，利用 f_c^2 和 f_c^4 的图像，证明 f_c 在 $c = -5/4$ 处经历了第二个倍周期分岔.
10. 找出一个分岔的例子，使得在分岔时会产生多于 3 个的不动点.
11. 证明 $f_3(x) = 3x(1-x)$ 在 I 上与 $f(x) = x^2 - 3/4$ 在 $\mathbb{R}$ 中的某个区间上是共轭的. 确定这个区间.
12. 假设 $f, g : [0,1] \to [0,1]$，并且存在一个从 f 到 g 的半共轭. 如果 f 是混沌的，证明 g 在 $[0,1]$ 上也是混沌的.
13. 证明 Σ 上的函数 $d(s,t)$ 满足距离函数的 3 个要求.
14. 确定如下给出的两个地址在三分康托集中所对应的点：
   (a) $LLRLLRLLR\cdots$;
   (b) $LRRLLRRLLRRL\cdots$.
15. 考虑帐篷映射

$$T(x) = \begin{cases} 2x & \text{如果 } 0 \leqslant x < 1/2, \\ -2x+2 & \text{如果 } 1/2 \leqslant x \leqslant 1. \end{cases}$$

   证明 T 在 $[0,1]$ 上是混沌的.
16. 考虑定义在整个 $\mathbb{R}$ 上的一个不同的"帐篷映射"

$$T(x) = \begin{cases} 3x & \text{如果 } x < 1/2, \\ -3x+3 & \text{如果 } 1/2 \leqslant x. \end{cases}$$

   确定轨道不趋于 $-\infty$ 的点构成的集合 Λ. 关于这个集合上的动力学你能说些什么?
17. 利用上题中的结果证明 15.5 节中的集合 Λ 同胚于三分康托集.
18. 证明下面的鞍结分岔定理：假设 f_λ 光滑地依赖于参数 λ，并且满足：
   (a) $f_{\lambda_0}(x_0) = x_0$;
   (b) $f'_{\lambda_0}(x_0) = 1$;
   (c) $f''_{\lambda_0}(x_0) \neq 0$;
   (d) $\left.\dfrac{\partial f_\lambda}{\partial \lambda}\right|_{\lambda=\lambda_0}(x_0) \neq 0$.

则存在关于 x_0 的一个区间 I 以及一个光滑函数 $\mu: I \to \mathbb{R}$, 使得 $\mu(x_0) = \lambda_0$, 并且满足

$$f_{\mu(x)}(x) = x.$$

进一步还有, $\mu'(x_0) = 0$ 以及 $\mu''(x_0) \neq 0$. 提示：对 $G(x, \lambda) = f_\lambda(x) - x$ 在 (x_0, λ_0) 处运用隐函数定理.

19. 讨论为什么鞍结分岔在只涉及一个不动点的分岔中是"典型的".

20. 回忆一下, 在第 14 章中, 理解洛伦茨系统的行为可以约化成理解某个定义在 $[-y^*, y^*]$ 上的一维函数 g 的动力学, 其中 g 的图像如图 15.14 所示. 注意, 对所有的 $y \neq 0$, 都有 $|g'(y)| > 1$, 而 g 在 0 处没有定义. 如图所示, 现在假设 $g^3(y^*) = 0$. 根据对称性, 我们还有 $g^3(-y^*) = 0$. 令 $I_0 = [-y^*, 0)$, $I_1 = (0, y^*]$, 并在 $[-y^*, y^*]$ 上定义通常的旅程映射.

    (a) 描述在 g 的迭代下所有可能的旅行路线构成的集合.

    (b) g 所有可能的周期点是什么？

    (c) 证明 g 在 $[-y^*, y^*]$ 上是混沌的.

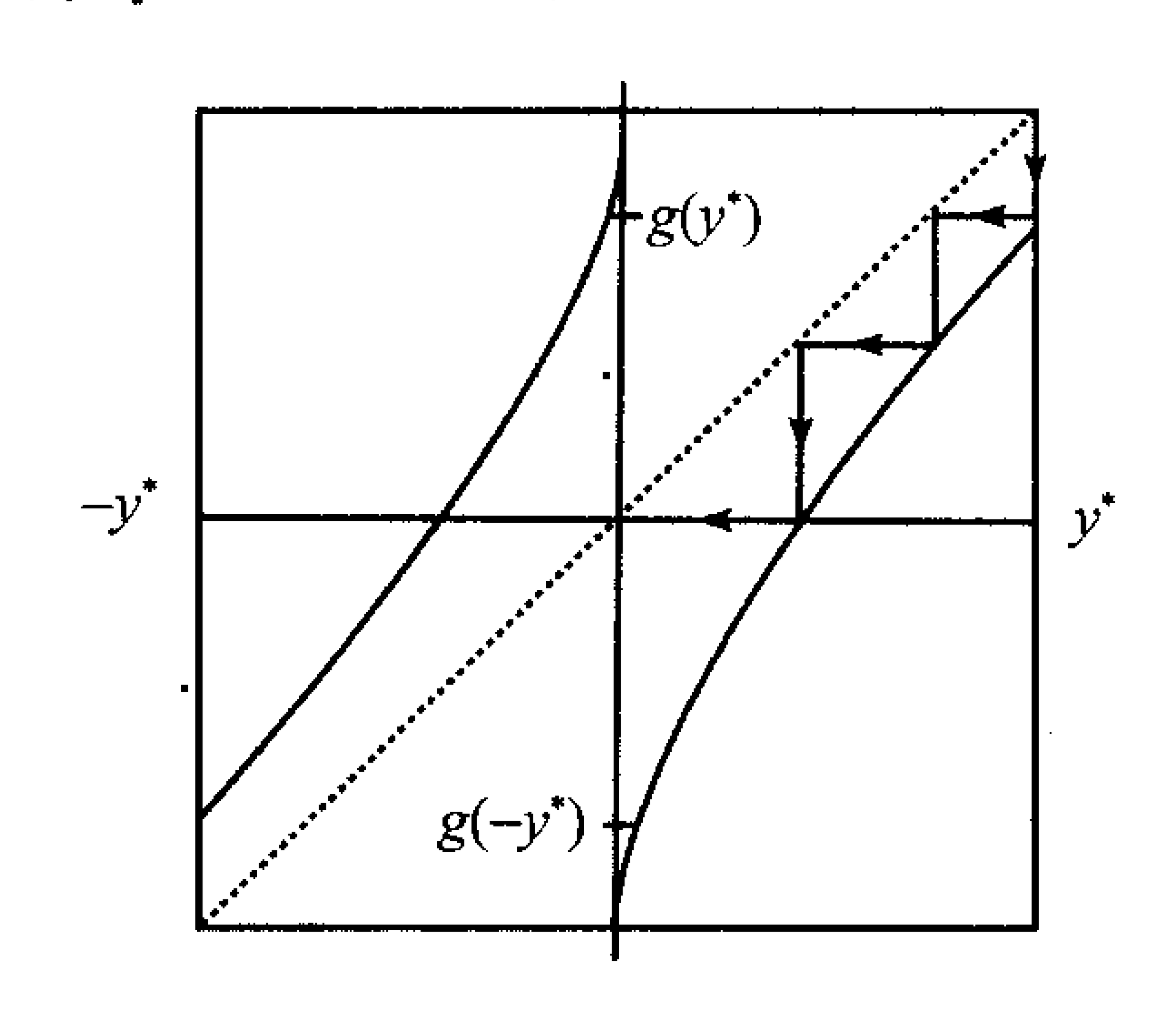


图 15.14　$[-y^*, y^*]$ 上一维函数 g 的图像

第 16 章　同宿现象

本章我们研究几个其他的三维微分方程系统，它们也都呈现出混沌行为. 这些系统包括 Shil'nikov 系统和双缠绕吸引子. 与处理洛伦茨系统一样，我们研究这些系统的主要手段就是先将它们约化成维数低一些的离散动力系统，然后再利用符号动力学来分析. 在这些情形，约化得到的离散系统是一个称为**马蹄映射**的平面映射. 它是首先被完整研究过的混沌系统之一.

16.1　Shil'nikov 系统

本节我们研究一个非线性微分方程系统的行为，它有一个螺线鞍点，并且这个平衡点有一个同宿解. 虽然我们在这里只是处理一个模型系统，但是 Shil'nikov 以及其他一些人的工作表明 [4, 40, 41]，这里所描述的现象在许多实际的微分方程系统中都存在. 事实上，在本章后面的实验中，我们将研究控制 Chua 电路的微分方程系统，对于某些参数值，这个系统就具有一对这样的同宿解.

在这个例子中，我们并不给出整个的微分方程系统，取而代之的是，我们首先在原点的某个圆柱形邻域中建立一个线性微分方程系统. 这个线性系统有一个二维的稳定曲面，其上的解都盘旋地趋于原点，它还有一个一维的不稳定曲线. 然后我们作一个简单但十分关键的动力学假设：两支不稳定曲线中的一支是一个同宿解，因而它最终会进入稳定曲面. 我们并不写出一个具有这种行为的具体微分方程. 虽然这样做是可能的，但是有了方程并不会对理解这个系统的大范围动力学有什么特别的帮助. 事实上，我们研究的这种现象只依赖于上面所描述的线性系统的定性性质，其中一个关键的不等式涉及该线性系统的本征值，以及我们的同宿假设.

这个系统的第一部分定义在 $\mathbb{R}^3$ 中的圆柱形区域 $\mathcal{S}$ 中，其中 $\mathcal{S}$ 由 $x^2+y^2 \leqslant 1$ 和 $|z| \leqslant 1$ 给出. 在这个区域中，考虑线性系统

$$\boldsymbol{X}' = \begin{pmatrix} -1 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \boldsymbol{X}.$$

相应的本征值为 $-1 \pm \mathrm{i}$ 和 2. 利用第 6 章的结果，我们可以很容易地得到这个系统的流 ϕ_t 为

$$x(t) = x_0 \mathrm{e}^{-t} \cos t + y_0 \mathrm{e}^{-t} \sin t,$$
$$y(t) = -x_0 \mathrm{e}^{-t} \sin t + y_0 \mathrm{e}^{-t} \cos t,$$
$$z(t) = z_0 \mathrm{e}^{2t}.$$

利用 xy 平面上的极坐标, $\mathcal{S}$ 中的解可以写得更简洁：

$$r(t) = r_0 \mathrm{e}^{-t},$$
$$\theta(t) = \theta_0 - t,$$
$$z(t) = z_0 \mathrm{e}^{2t}.$$

这个系统具有二维的稳定平面 (xy 平面) 和一对分别位于 z 轴的正负半轴上的不稳定曲线 $\zeta^{\pm}$.

需要注意的是, 我们在这个系统中选取的本征值没有什么特别的含义. 事实上, 对于任何的本征值 $\alpha \pm \mathrm{i}\beta$ 和 λ, 其中 $\alpha < 0, \beta \neq 0, \lambda > 0$, 只要它们满足 $\lambda > -\alpha$ 这个重要条件, 下面的每件事都一样地成立.

$\mathcal{S}$ 的边界由三部分构成：上下两个圆盘 $D^{\pm}$, 它们分别由 $z = \pm 1, r \leqslant 1$ 给出, 以及圆柱形边界 C, 它由 $r = 1, |z| \leqslant 1$ 给出. 稳定平面与 C 交于圆周 $z = 0$, 而且它将 C 分成上下两半, C^+ 和 C^-, 它们分别对应于 $z > 0$ 和 $z < 0$. 我们可以将 $D^{\pm}$ 用 r 和 θ 参数化, 而在 C 上则选取参数 θ 和 z. 在本节, 我们的注意力集中在 C^+ 上.

这个系统从 C^+ 出发的任何解最终都要从 D^+ 离开 $\mathcal{S}$. 如果跟随从 C^+ 出发的解曲线一直到与 D^+ 相交, 我们就可以定义一个映射 $\psi_1 : C^+ \to D^+$. 给定 $(\theta_0, z_0) \in C^+$, 记 $\tau = \tau(\theta_0, z_0)$ 为经过 (θ_0, z_0) 的解到达 D^+ 所花的时间. 利用 $z(t) = z_0 \mathrm{e}^{2t}$ 可以立刻算得 $\tau = -\ln\sqrt{z_0}$. 因而,

$$\psi_1 \begin{pmatrix} 1 \\ \theta_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} r_1 \\ \theta_1 \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{z_0} \\ \theta_0 + \ln\sqrt{z_0} \\ 1 \end{pmatrix}.$$

为了简单起见, 我们将 ψ_1 看成是一个从圆柱面 (θ_0, z_0) 到平面 (r_1, θ_1) 上的映射. 注意, C^+ 中的竖直线段 $\theta_0 = \theta^*$ 被 ψ_1 映到了螺旋线

$$z_0 \to \left(\sqrt{z_0}, \theta^* + \ln\sqrt{z_0}\right),$$

由于当 $z_0 \to 0$ 时, $\ln\sqrt{z_0} \to -\infty$, 这条螺旋线将盘旋地趋于 D^+ 中的点 $r = 0$.

为了定义这个系统的第二部分, 我们假设不稳定曲线上从 D^+ 离开原点的那一支, 即 ζ^+ 是一个同宿解. 即, ζ^+ 最终要回到稳定平面上 (见图 16.1). 我们假设

ζ^+ 第一次与圆柱面 C 交于点 $r=1, \theta=0, z=0$. 更准确些, 我们假设存在一个时刻 t_1, 使得在 r, θ, z 坐标下有 $\phi_{t_1}(0,\theta,1)=(1,0,0)$.

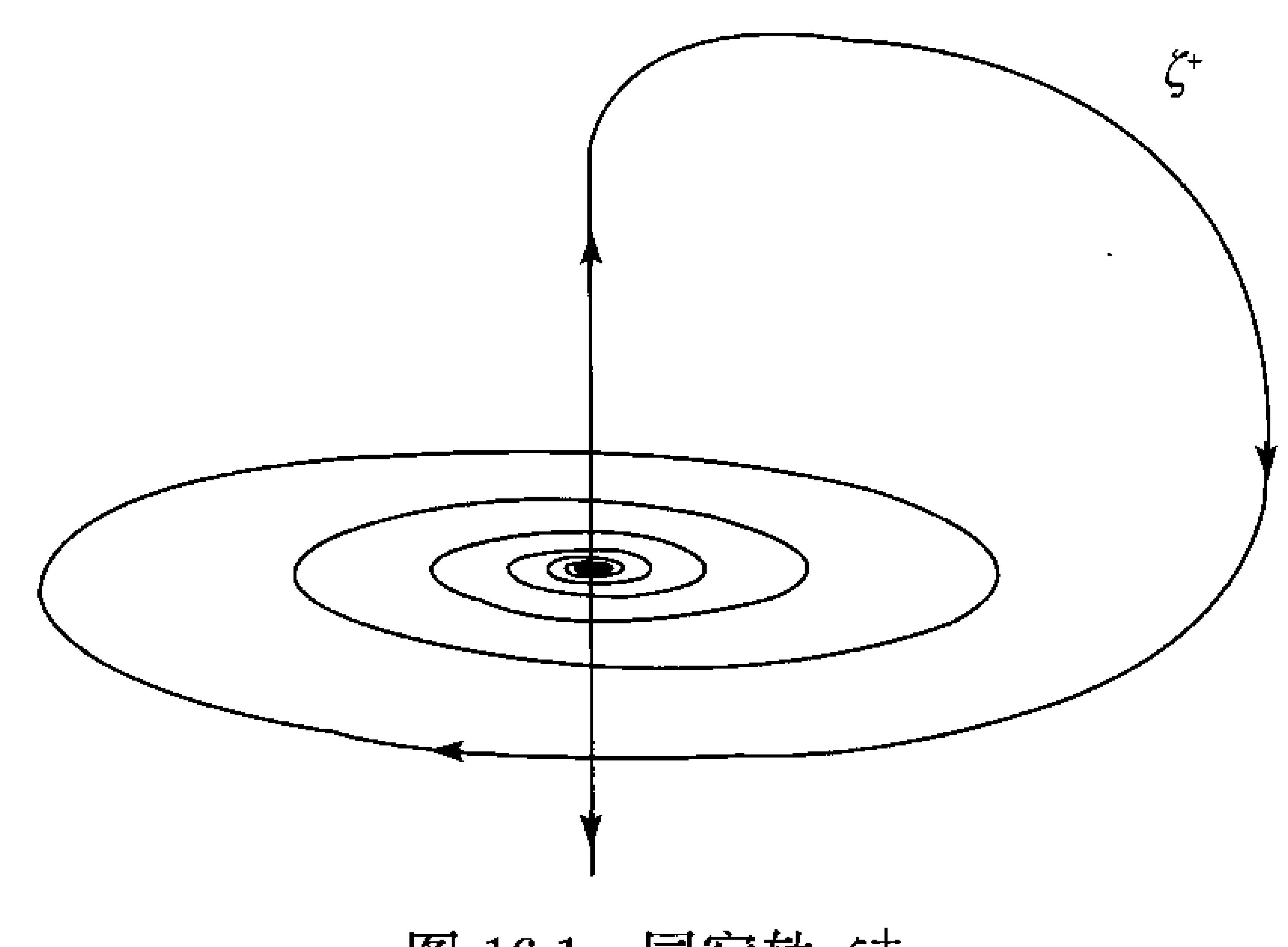


图 16.1　同宿轨 ζ^+

于是, 跟随从 D^+ 上 $r=0$ 附近出发的解一直到与 C 相交, 我们就可以定义第二个映射 ψ_2. 我们假设 ψ_2 事实上在整个 D^+ 上都可定义. 在 D^+ 上的笛卡儿坐标下, 我们假设 ψ_2 按如下的方式将 $(x,y)\in D^+$ 带到 $(\theta_1, z_1)\in C$:

$$\psi_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \theta_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} y/2 \\ x/2 \end{pmatrix}.$$

在极坐标下, ψ_2 的表达式为

$$\begin{aligned} \theta_1 &= (r\sin\theta)/2, \\ z_1 &= (r\cos\theta)/2. \end{aligned}$$

当然, 这是一个主要的假设, 因为要想对一个特定的非线性系统写出这个映射事实上是不可能的.

现在, 复合函数 $\Phi=\psi_2\circ\psi_1$ 就定义了 C^+ 上的一个庞加莱映射. 映射 ψ_1 定义在 C^+ 上, 取值于 D^+ 中, 然后, ψ_2 取值于 C 中. 我们有 $\Phi: C^+\to C$, 其中

$$\Phi\begin{pmatrix} \theta_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} \theta_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\sqrt{z_0}\sin\left(\theta_0+\ln\sqrt{z_0}\right) \\ \frac{1}{2}\sqrt{z_0}\cos\left(\theta_0+\ln\sqrt{z_0}\right) \end{pmatrix}.$$

见图 16.2.

与洛伦茨系统中一样，现在我们将研究这个三维系统的流约化成了研究一个平面离散动力系统. 在下节我们将看到，这种类型映射的动力学是惊人的丰富，而且可以 (部分地) 利用符号动力学来分析. 为了能稍稍领略一下随后的内容，我们来寻找 Φ 的不动点. 为此，我们需要求解方程：

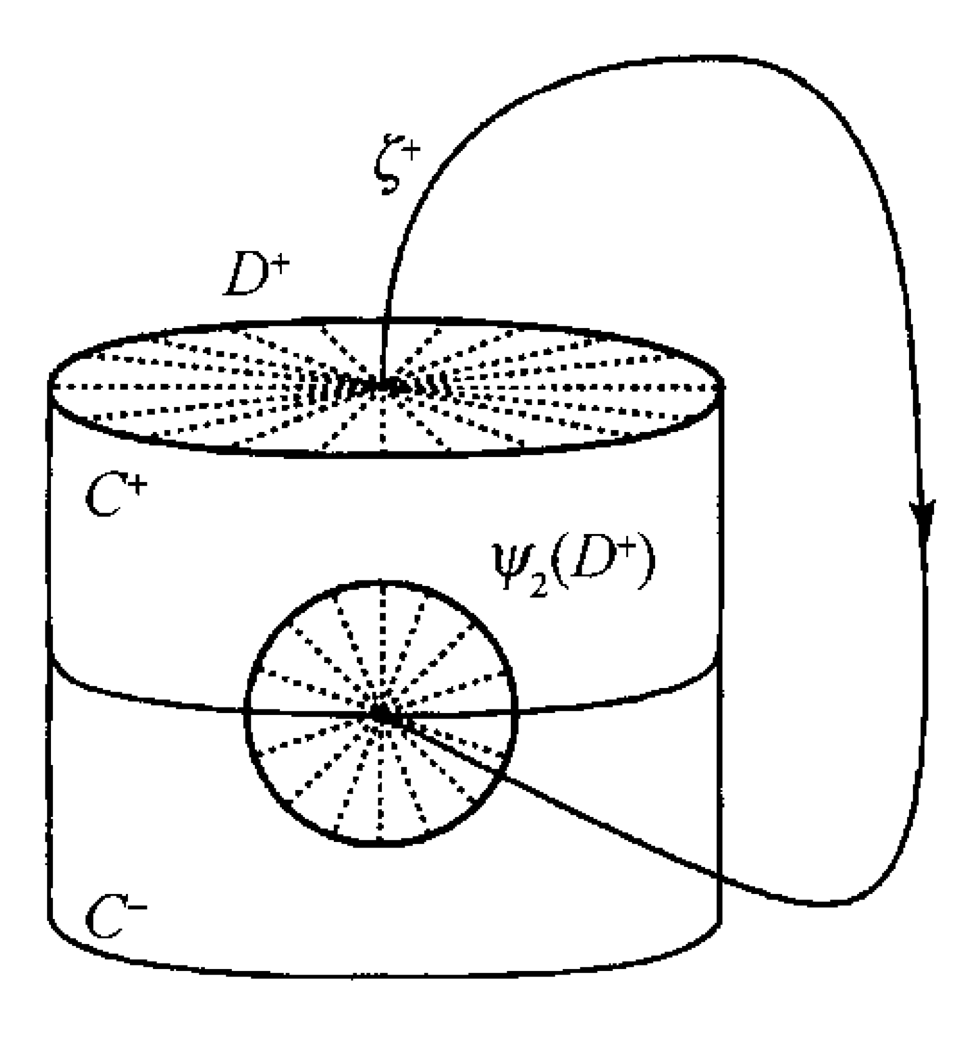


图 16.2 映射 $\psi_2 : D^+ \to C$

$$\theta_0 = \frac{1}{2}\sqrt{z_0}\sin\left(\theta_0 + \ln\sqrt{z_0}\right)$$
$$z_0 = \frac{1}{2}\sqrt{z_0}\cos\left(\theta_0 + \ln\sqrt{z_0}\right).$$

这些方程看起来相当可怕. 但是，将两个方程平方之后再相加，我们得到

$$\theta_0^2 + z_0^2 = \frac{z_0}{4},$$

从而，

$$\theta_0 = \pm\frac{1}{2}\sqrt{z_0 - 4z_0^2},$$

它在 $0 \leqslant z_0 \leqslant 1/4$ 时是有定义的. 将这个表达式代入上面的第二个方程，我们得到如下的待解方程：

$$\cos\left(\pm\frac{1}{2}\sqrt{z_0 - 4z_0^2} + \ln\sqrt{z_0}\right) = 2\sqrt{z_0}.$$

现在，当 $z_0 \to 0$ 时，$\sqrt{z_0 - 4z_0^2}$ 这一项趋于零，而 $\ln\sqrt{z_0} \to -\infty$. 因而，当 $z_0 \to 0$ 时，上述方程左端的图像将在 ± 1 之间振动无穷多次. 于是，这个图像就要与 $2\sqrt{z_0}$ 的图像相交无穷多次，因此这个方程有无穷多个解. 相应地，Φ 就有无穷多个不动点. 每个这种不动点都对应于这个系统从 C^+ 出发的一个周期解，这种周期解先在原点附近绕 z 轴缠绕很多次，然后沿同宿轨附近前进直到其回到 C^+ 中封闭起来. 见图 16.3.

现在我们来描述这个映射的几何. 在下节，我们将利用这些想法去研究这个映射一个简化形式的动力学. 首先注意，C^+ 中的圆周 $z_0 = \alpha$ 被 ψ_1 映到 D^+ 中的中心在 $r = 0$ 的圆周 $r = \sqrt{\alpha}$，这是因为

$$\psi_1\begin{pmatrix}\theta_0\\ \alpha\end{pmatrix} = \begin{pmatrix}r_1\\ \theta_1\end{pmatrix} = \begin{pmatrix}\sqrt{\alpha}\\ \theta_0 + \ln\sqrt{\alpha}\end{pmatrix}.$$

然后，ψ_2 将这些圆周映到 C 中中心在 $\theta_1 = z_1 = 0$、半径为 $\sqrt{\alpha}/2$ 的圆周. (严格地讲，这些是 θz 平面上的圆周，在圆柱面上这些圆周被“弯曲”了.) 特别地，我们看到定义域 C^+ 的一半被映到了圆柱面的下半部分 C^-，因而将不再起作用.

记 H 为半圆盘 $\Phi(C^+)\cap\{z\geqslant 0\}$, 它的中心在 $\theta_1=z_1=0$, 半径为 1/2. H 在 C^+ 中的原像由所有这样的点 (θ_0,z_0) 构成, 其中 (θ_0,z_0) 的像满足 $z_1\geqslant 0$, 因而我们必定有

$$z_1=\frac{1}{2}\sqrt{z_0}\cos\left(\theta_0+\ln\sqrt{z_0}\right)\geqslant 0.$$

由此可知 H 的原像为

$$\Phi^{-1}(H)=\{(\theta_0,z_0)|-\pi/2\leqslant\theta_0+\ln\sqrt{z_0}\leqslant\pi/2\},$$

其中 $0<z_0\leqslant 1$. 这是由两条曲线 $\theta_0+\ln\sqrt{z_0}=\pm\pi/2$ 所界定的区域, 其中的每条曲线都在 C^+ 中盘旋而下地趋于圆周 $z=0$. 见图 16.4. 这是因为, 当 $z_0\to 0$ 时, 必定有 $\theta_0\to\infty$. 更一般地, 对于 $-\pi/2\leqslant\alpha\leqslant\pi/2$, 考虑曲线

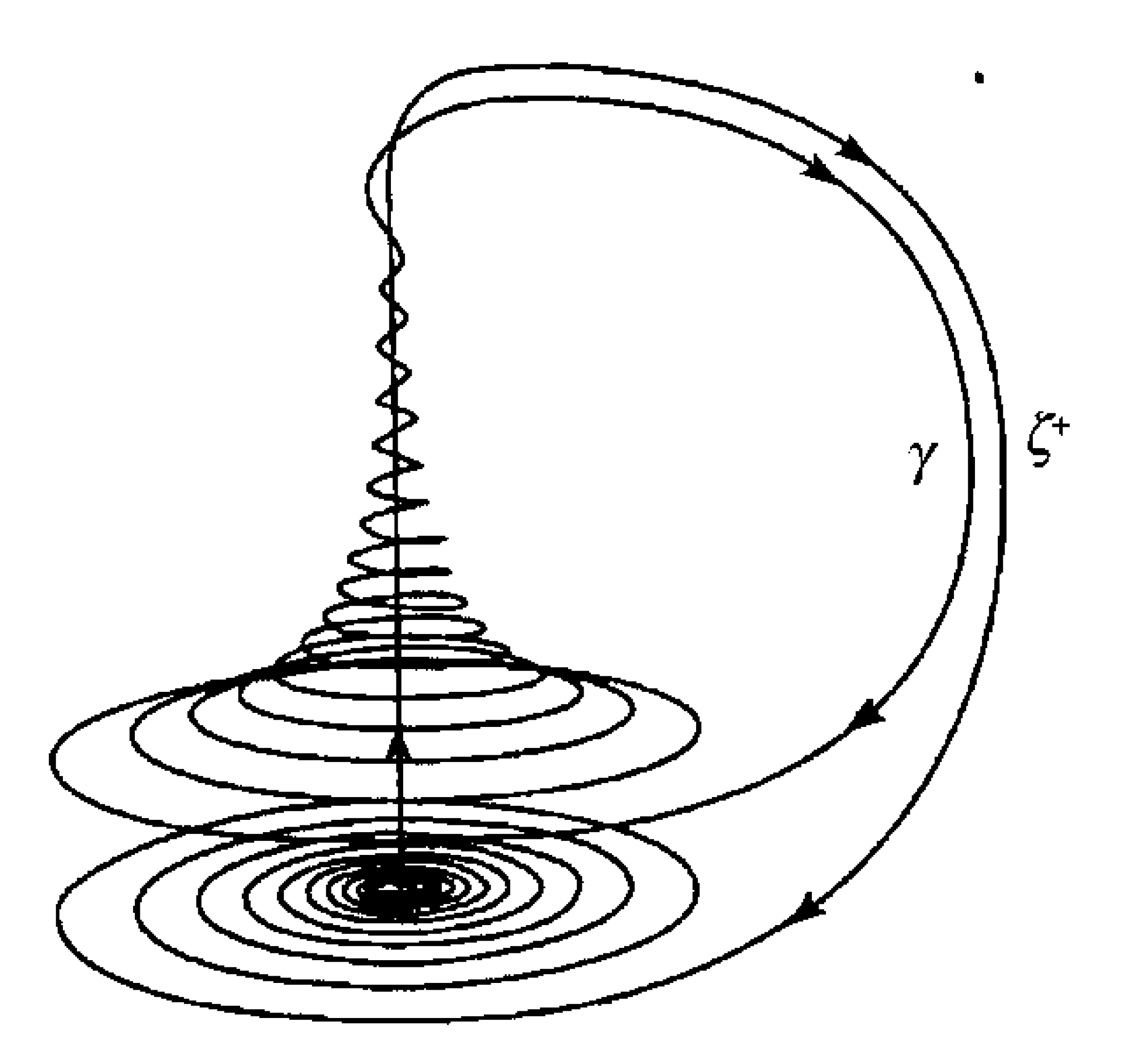


图16.3　在同宿解 ξ^+ 附近的一个周期解 γ

$$l_\alpha:\quad \theta_0+\ln\sqrt{z_0}=\alpha.$$

所有这些曲线充满了原像 $\Phi^{-1}(H)$, 并且每条这样的曲线都像原像的边界曲线一样绕着 C 盘旋. 现在我们有

$$\Phi(l_\alpha)=\frac{\sqrt{z_0}}{2}\begin{pmatrix}\sin\alpha\\ \cos\alpha\end{pmatrix},$$

因而, Φ 将每个 l_α 映到了 C^+ 中一条从 $\theta=z=0$ 出发的射线, 并且这些射线可由 $\sqrt{z_0}$ 参数化. 特别地, Φ 将每条边界曲线 $l_{\pm\pi/2}$ 映到了 C 中的 $z=0$.

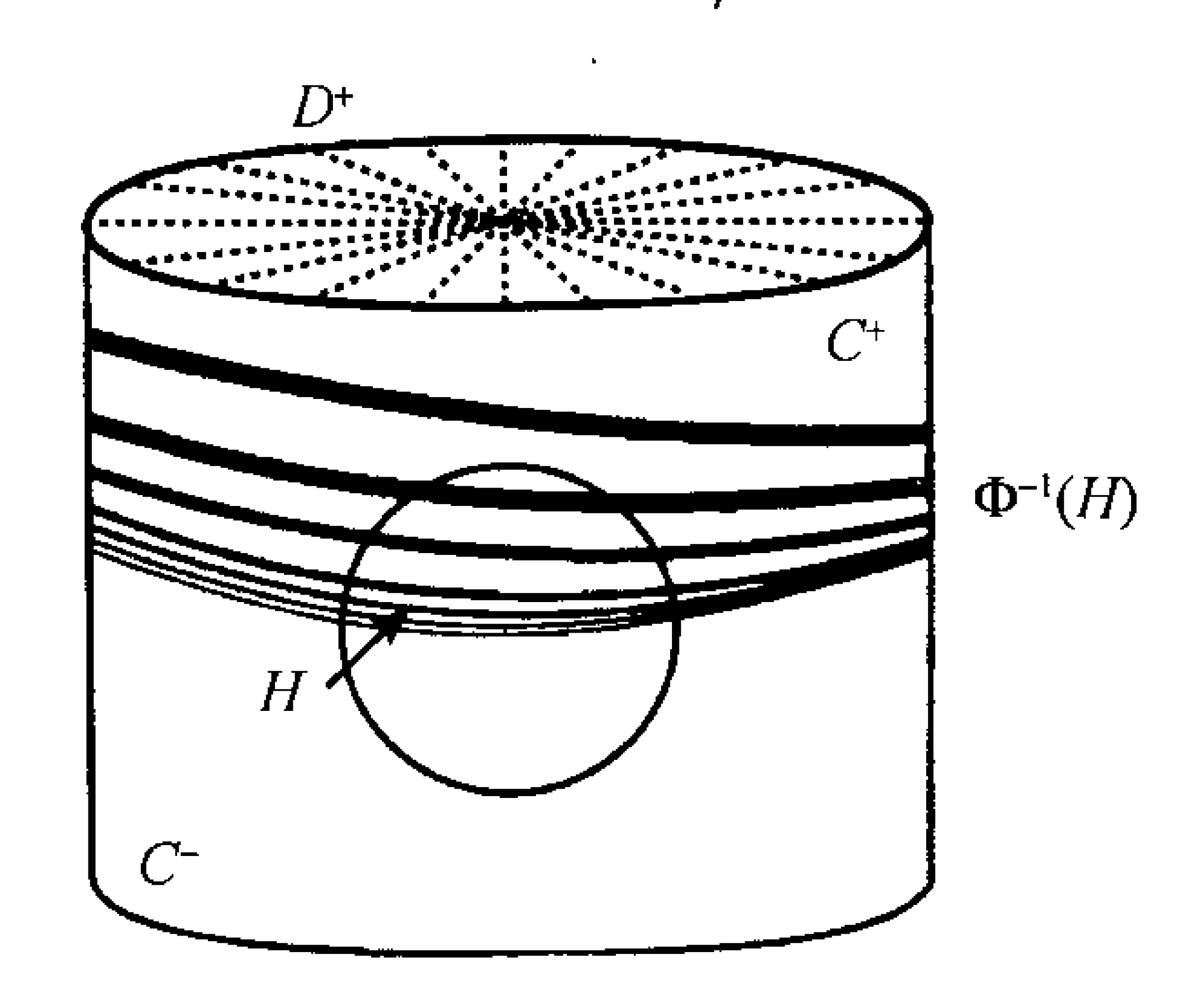


图 16.4　半圆盘 H 以及它在 C^+ 中的原像

由于曲线 $l_{\pm\pi/2}$ 盘旋而下地趋于 C 中的圆周 $z=0$, 因而 $\Phi^{-1}(H)$ 和 H 相交于无穷多个窄条, 这些窄条在接近 $z=0$ 时是几乎水平的 (见图 16.4). 我们将这些

窄条记为 H_k, 其中 k 充分大. 准确地说, 记 H_k 为 $\Phi^{-1}(H)\cap H$ 中满足

$$2k\pi-\frac{1}{2}\leqslant\theta_0\leqslant 2k\pi+\frac{1}{2}$$

的连通分支. H_k 的上边界为螺线 $l_{\pi/2}$ 的一部分, 而 H_k 的下边界为螺线 $l_{-\pi/2}$ 的一部分. 利用

$$-\frac{\pi}{2}\leqslant\theta_0+\ln\sqrt{z_0}\leqslant\frac{\pi}{2}$$

这个事实, 我们可得, 如果 $(\theta_0,z_0)\in H_k$, 则有

$$-(4k+1)\pi-1\leqslant-\pi-2\theta_0\leqslant 2\ln\sqrt{z_0}\leqslant\pi-2\theta_0\leqslant-(4k-1)\pi+1,$$

由此可得

$$\exp(-(4k+1)\pi-1)\leqslant z_0\leqslant\exp(-(4k-1)\pi+1).$$

现在考虑 H_k 在 Φ 下的像. H_k 的上下边界都被映到 $z=0$. 曲线 $l_\alpha\cap H_k$ 被映到了从 $\theta=z=0$ 出发的射线上的弧段. 这些射线就是上面给出的

$$\frac{\sqrt{z_0}}{2}\begin{pmatrix}\sin\alpha\\ \cos\alpha\end{pmatrix}.$$

特别地, 曲线 l_0 被映到了竖直线 $\theta_1=0, z_1=\sqrt{z_0}/2$ 上. 利用上面对 H_k 中 z_0 的大小的估计, 容易验证, 当 $k\geqslant 2$ 时, l_0 的像全部位于 H_k 的上方. 于是, 如图 16.5 所示, 像集 $\Phi(H_k)$ 是一个 "马蹄形" 的区域, 并且两次穿过 H_k. 特别地, 当 k 很大时, 曲线 $l_\alpha\cap H_k$ 与 $\Phi(H_k)$ 交于两个几乎水平的子弧段.

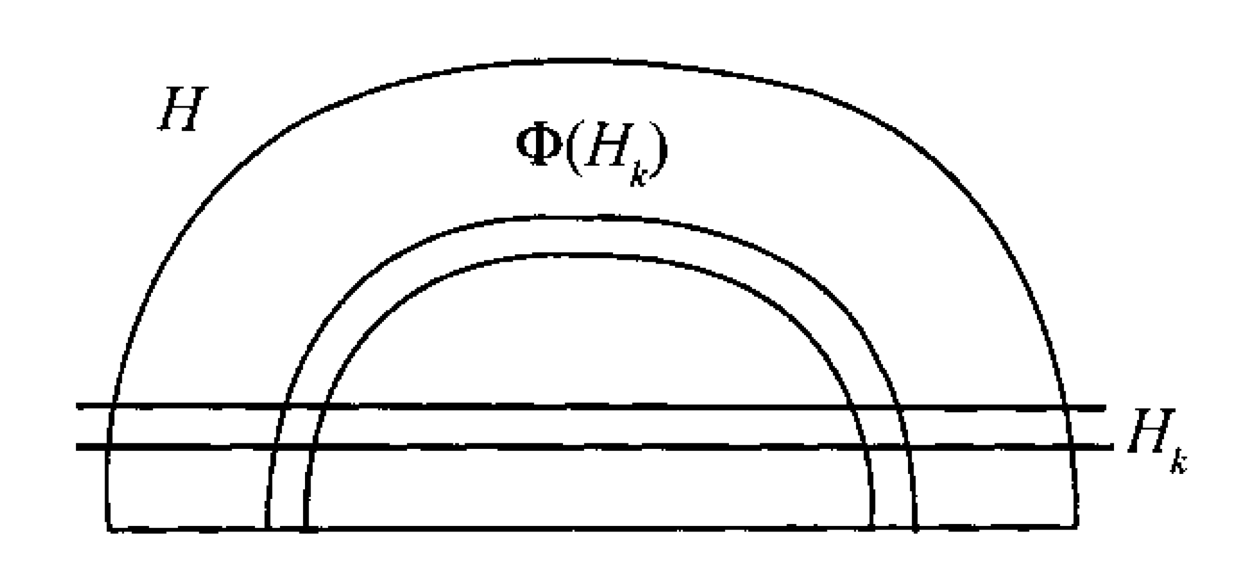


图 16.5 H_k 的像是一个马蹄并且两次穿过 H_k

这样的映射称为**马蹄映射**. 在下节, 我们将讨论这类映射的原型.

16.2 马蹄映射

符号动力学不仅对于我们理解一维合理映射起到了关键的作用, 而且还可以用来研究高维现象. 在本节, 我们来描述 $\mathbb{R}^2$ 中的一个重要例子 —— 马蹄映射 [43]. 我们将看到, 这个映射与上节中描述的庞加莱映射有许多的共同之处.

为了定义马蹄映射, 我们考虑一个区域 D, 它由三个部分组成: 中间的正方形 S, 其边长为 1, 再加上上下两个半圆盘 D_1 和 D_2. D 的形状像一个"体育场".

马蹄映射 F 将 D 映入自己的内部, 其具体定义描述如下: 首先, F 将 S 沿水平方向线性地压缩到原来的一个因子 $\delta < 1/2$, 并且将 S 沿竖直方向线性地扩张到原来的 $1/\delta$ 倍, 从而 S 经过作用之后就变成一个又高又窄的矩形, 然后 F 再将这个矩形弯曲后放回到 D 内部, 从而形成一个如图 16.6 所示的马蹄状图形. 我们规定, F 将 S 线性地映到马蹄的两条竖直"腿"上.

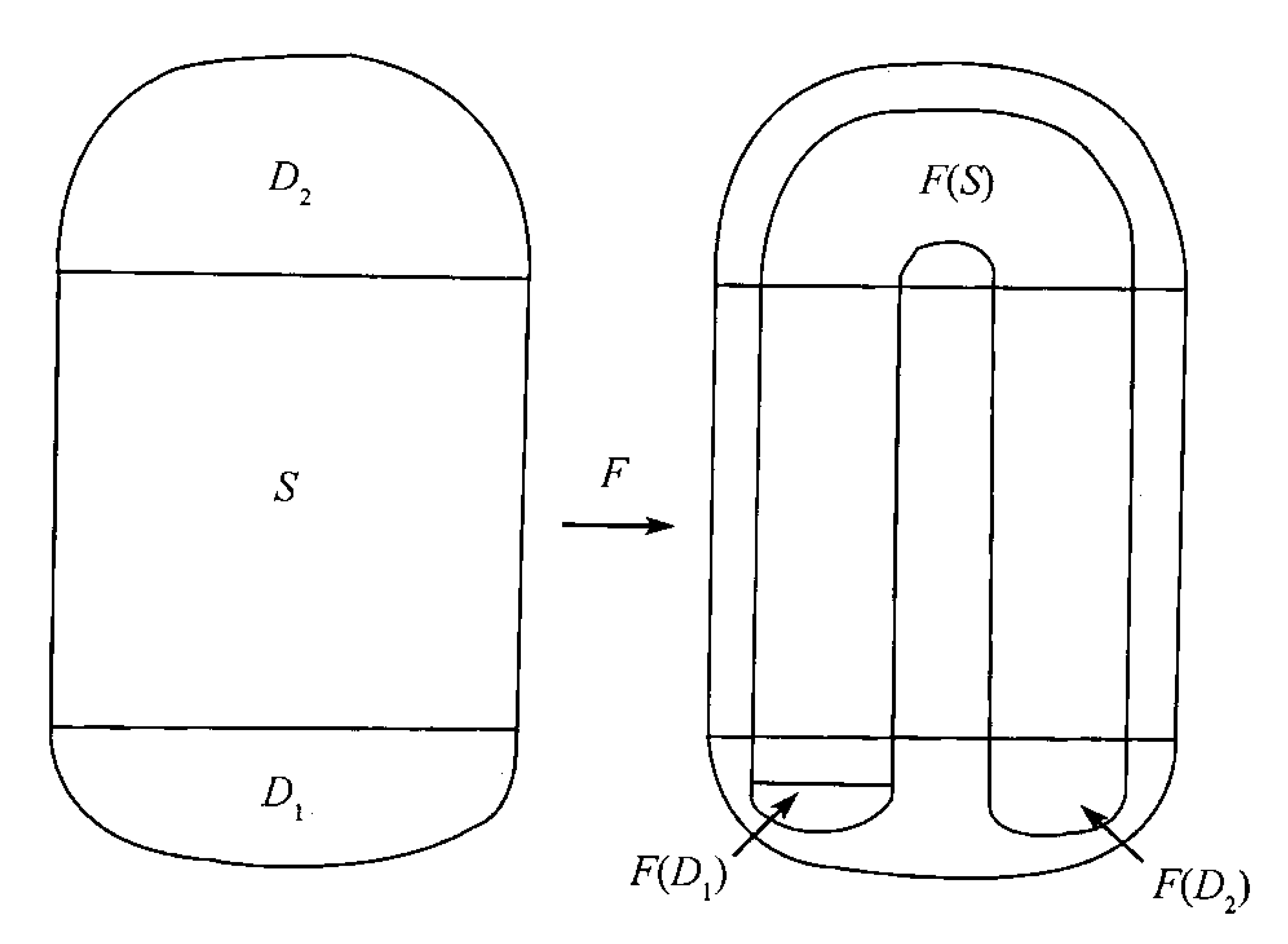


图 16.6 马蹄映射的第一次迭代

我们假定两个半圆盘区域 D_1 和 D_2 如图中那样被映入 D_1 的内部. 我们还假定 D_1 中有一个不动点, 并且它要吸引 D_1 中所有的其他轨道. 注意, $F(D) \subset D$, 并且 F 是一对一的. 但是, 因为 F 不是满的, 因而 F 的逆并不是总有定义. 本节后面的部分就是研究 F 在 D 中的动力学.

首先注意, S 的原像由两个水平矩形 H_0 和 H_1 构成, 这两个水平矩形 H_0 和 H_1 分别被线性地映满 $F(S) \cap S$ 的两个竖直分支 V_0 和 V_1. 从而, V_0 和 V_1 的宽度都是 δ, 而 H_0 和 H_1 的高度也都是 δ(见图 16.7). 由于 $F: H_0 \to V_0$ 以及 $F: H_1 \to V_1$ 都是线性的, 因而 F 将 H_j 中的水平线段和竖直线段分别映到 V_j 中的水平线段和竖直线段, 这里 $j = 1, 2$. 由此可以推出, 如果 h 和 $F(h)$ 都是 S 中的水平线段, 则 $F(h)$ 的长度等于 h 的长度乘以 δ. 类似地, 如果 v 是 S 中的一条竖直线段, 并且它的像 $F(v)$ 也在 S 中, 则 $F(v)$ 的长度等于 v 的长度乘以 $1/\delta$.

我们现在来描述 D 中任意一点 $\boldsymbol{X}$ 的**前向轨道**. 根据定义, $\boldsymbol{X}$ 的前向轨道由 $\{F^n(\boldsymbol{X}) | n \geqslant 0\}$ 给出. 根据前面的假设, F 在 D_1 中有唯一的不动点 $\boldsymbol{X}_0$, 并且对于所有的 $\boldsymbol{X} \in D_1$, 都有 $\lim_{n\to\infty} F^n(\boldsymbol{X}) = \boldsymbol{X}_0$. 又由于 $F(D_2) \subset D_1$, 因而 D_2 中所有前向轨道的行为也是如此. 类似地, 如果 $\boldsymbol{X} \in S$, 但对某个 $k > 0$ 有 $F^k(\boldsymbol{X}) \notin S$, 则

必有 $F^k(\boldsymbol{X}) \in D_1 \cup D_2$, 因而当 $n \to \infty$ 时, 同样有 $F^n(\boldsymbol{X}) \to \boldsymbol{X}_0$. 这样, 对于 D 中的任意一点 $\boldsymbol{X}$, 只要它的轨道进入 D_1, 我们就了解了它的前向轨道. 从而, 我们只需考虑那些前向轨道从不进入 D_1 的点, 当然这些点的前向轨道整个地位于 S 中. 令

$$\Lambda_+ = \{\boldsymbol{X} \in S | F^n(\boldsymbol{X}) \in S,\ n = 0, 1, 2, \cdots\}.$$

我们断言, Λ_+ 的性质与第 15 章描述的一维合理映射的相应集合类似.

如果 $\boldsymbol{X} \in \Lambda_+$, 则 $F(\boldsymbol{X}) \in S$, 因而必定有 $\boldsymbol{X} \in H_0$ 或 $\boldsymbol{X} \in H_1$, 这是因为 S 中的所有其他的点都被映入 D_1 或 D_2 中. 由于 $F^2(\boldsymbol{X})$ 也在 S 中, 因而我们还有 $F(\boldsymbol{X}) \in H_0 \cup H_1$, 于是 $\boldsymbol{X} \in F^{-1}(H_0 \cup H_1)$. 这里 $F^{-1}(W)$ 代表集合 W 在 D 中的原像. 一般地, 由于 $F^n(\boldsymbol{X}) \in S$, 因而有 $\boldsymbol{X} \in F^{-n}(H_0 \cup H_1)$. 于是可以将 Λ_+ 写成

$$\Lambda_+ = \bigcap_{n=0}^{\infty} F^{-n}(H_0 \cup H_1).$$

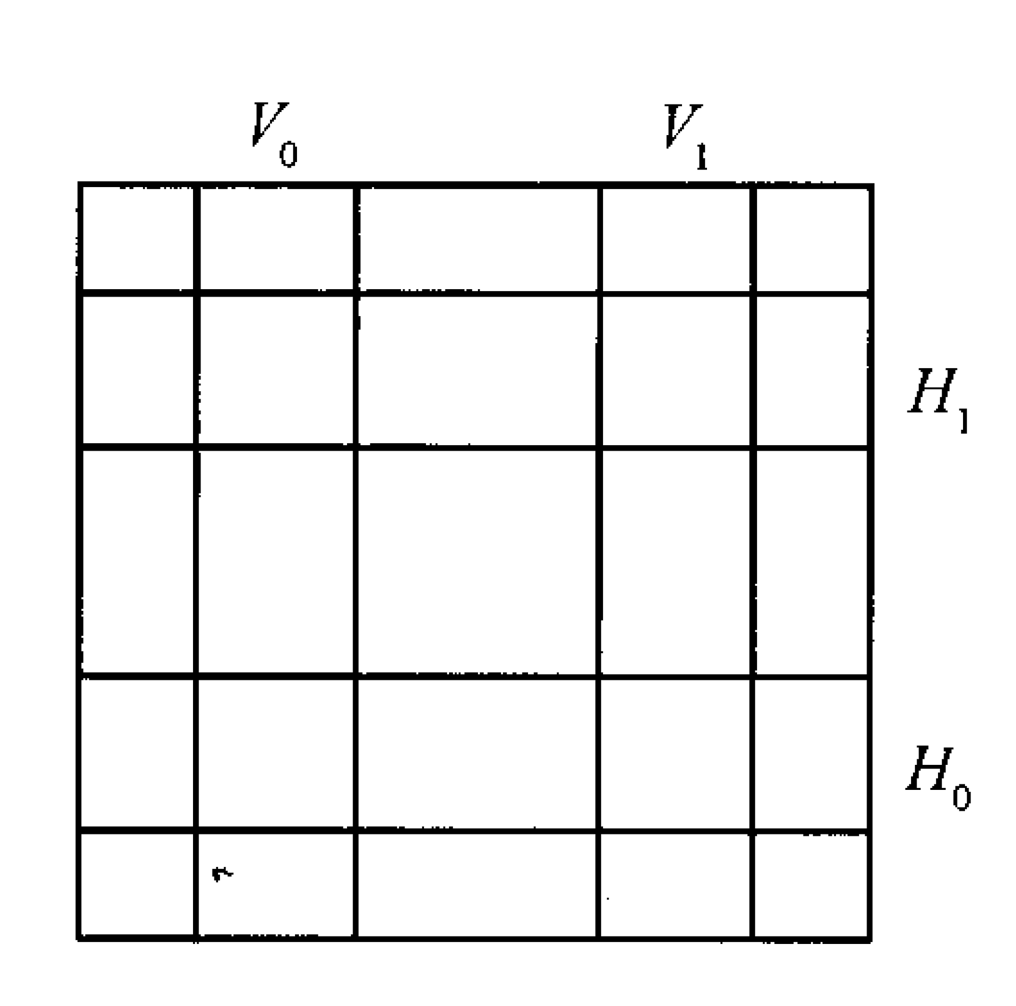


图 16.7 矩形 H_0 和 H_1 以及它们的像 V_0 和 V_1

现在, 如果 H 是连接 S 左右两边、高度为 h 的一个水平条, 则 $F^{-1}(H)$ 就由一对更窄的、高度为 δh 的水平条构成, 并且这两个更窄的水平条分别位于 H_0 和 H_1 中. 而且这两个更窄的水平条在 F 下的像分别为 $H \cap V_0$ 和 $H \cap V_1$. 特别地, 如果 $H = H_i$, 则 $F^{-1}(H_i)$ 是一对水平条, 高度都是 δ^2, 且一个在 H_0 中, 另一个在 H_1 中. 类似地, $F^{-1}(F^{-1}(H_i)) = F^{-2}(H_i)$ 由 4 个水平条构成, 每一个的高度都是 δ^3, 一般地, $F^{-n}(H_i)$ 由 2^n 个高度都是 δ^{n+1} 的水平条构成. 于是, 运用 15.5 节中同样的方法, 可以证明 Λ_+ 是一些线段的"康托集", 其中的每条线段都水平地贯穿 S.

马蹄映射与合理映射之间的主要区别在于, 对于马蹄映射, 后向轨道是唯一的, 而对于合理映射, 后向轨道则可能有无穷多条. 给定 $\boldsymbol{X} \in S$, 如果对所有的 $n \geqslant 1$, $F^{-n}(\boldsymbol{X})$ 都在 D 中有定义, 则 $\boldsymbol{X}$ 的**后向轨道**就是 $\{F^{-n}(\boldsymbol{X}) | n = 1, 2, \cdots\}$. 如果某个 $F^{-n}(\boldsymbol{X})$ 没有定义, 则 $\boldsymbol{X}$ 的后向轨道终止. 记 Λ_- 是由后向轨道对所有 n 都有定义的点构成的集合, 显然, Λ_- 中点的整条后向轨道都位于 S 中. 如果 $\boldsymbol{X} \in \Lambda_-$, 则对所有的 $n \geqslant 1$, 都有 $F^{-n}(\boldsymbol{X}) \in S$, 即 $\boldsymbol{X} \in F^n(S)$. 与上面一样, 这就迫使对所有的 $n \geqslant 1$, 都有 $\boldsymbol{X} \in F^n(H_0 \cup H_1)$. 因而我们可以将 Λ_- 写成

$$\Lambda_- = \bigcap_{n=1}^{\infty} F^n(H_0 \cup H_1).$$

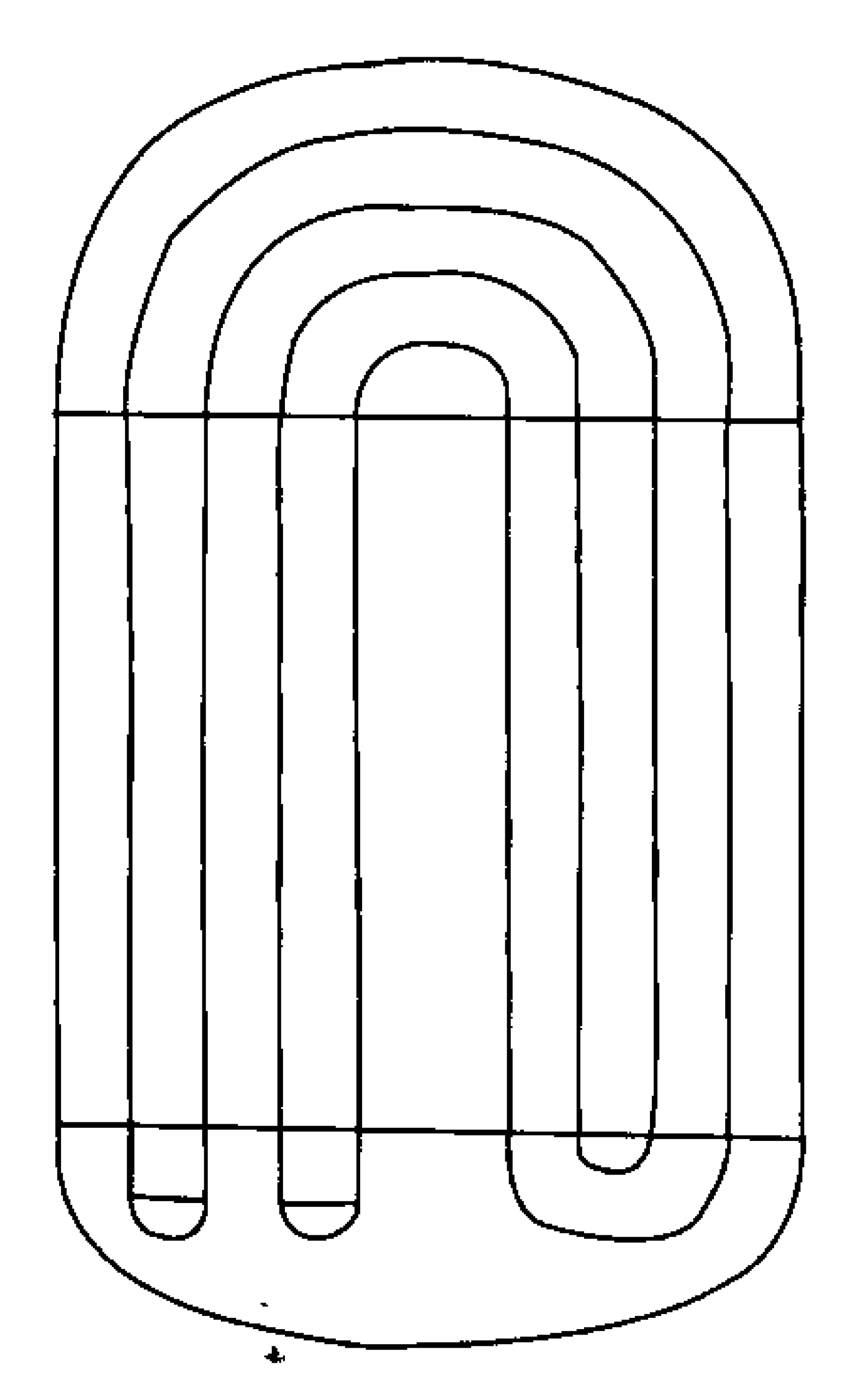


图 16.8 马蹄映射的第二次迭代

另一方面, 如果有 $\boldsymbol{X} \in S$ 和 $F^{-1}(\boldsymbol{X}) \in S$, 则必定有 $\boldsymbol{X} \in F(S) \cap S$, 因而 $\boldsymbol{X} \in V_0$ 或 $\boldsymbol{X} \in V_1$. 类似地, 如果还有 $F^{-2} \in S$, 则 $\boldsymbol{X} \in F^2(S) \cap S$, 其中 $F^2(S) \cap S$ 由 4 个更窄的竖条构成, 2 个在 V_0 中, 2 个在 V_1 中. 在图 16.8 中, 我们画出了 D 在 F^2 下的像. 通过完全类似于上面的讨论, 容易验证 Λ_- 是一些竖直线段的"康托集".

令

$$\Lambda = \Lambda_+ \cap \Lambda_-$$

为这两个集合的交集. Λ 中任一点的整条轨道 (后向轨道和前向轨道) 都位于 S 中.

为了在现在的图像中引入符号动力学, 我们将 Λ 中的每个点赋以一个双边无穷的 0, 1 序列. 如果 $\boldsymbol{X} \in \Lambda$, 则从上面的讨论可知,

$$\boldsymbol{X} \in \bigcap_{n=-\infty}^{\infty} F^n(H_0 \cup H_1).$$

于是我们将 $\boldsymbol{X}$ 关联到它的**旅程**

$$S(\boldsymbol{X}) = (\cdots s_{-2}s_{-1} \cdot s_0 s_1 s_2 \cdots),$$

其中 $s_j = 0$ 或 1, 并且 $s_j = k$ 当且仅当 $F^j(\boldsymbol{X}) \in H_k$. 这就产生了 Λ 上的符号动力学. 记 Σ_2 为所有双边无穷的 0, 1 序列构成的集合:

$$\Sigma_2 = \{(\mathbf{s}) = (\cdots s_{-2}s_{-1} \cdot s_0 s_1 s_2 \cdots) | s_j = 0\text{或}1\}.$$

与 15.5 节一样, 我们在 Σ_2 上定义一个如下的距离函数:

$$d((\mathbf{s}), (\mathbf{t})) = \sum_{i=-\infty}^{\infty} \frac{|s_i - t_i|}{2^{|i|}}.$$

于是, Σ_2 中的两个序列是"接近的", 当且仅当对于某个 (很大的)n, 这两个序列在所有满足 $|k| \leqslant n$ 的第 k 个位置上都相同. 我们定义 (双边)**移位映射**σ 如下:

$$\sigma(\cdots s_{-2}s_{-1} \cdot s_0 s_1 s_2 \cdots) = (\cdots s_{-2}s_{-1}s_0 \cdot s_1 s_2 \cdots).$$

即, σ 的作用就是简单地将 Σ_2 中的每个序列向左移动一个位置 (等价地, σ 将小数点向右移动一个位置). 与前面的 (单边) 移位映射不同, 这个 (双边) 移位映射是有

逆的. 显然, 向右移动一个位置就给出了它的逆. 容易验证 σ 是 Σ_2 上的一个同胚 (见本章的习题 2).

现在, 移位映射就是限制在 Λ 上 F 的模型. 事实上, 旅程映射 S 给出了 Λ 上的 F 和 Σ_2 上的 σ 之间的一个共轭. 理由如下: 如果 $\boldsymbol{X} \in \Lambda$, 且 $S(\boldsymbol{X}) = (\cdots s_{-2}s_{-1} \cdot s_0 s_1 s_2 \cdots)$, 则有 $\boldsymbol{X} \in H_{s_0}, F(\boldsymbol{X}) \in H_{s_1}, F^{-1}(\boldsymbol{X}) \in H_{s_{-1}}$, 等等. 这样就有 $F(\boldsymbol{X}) \in H_{s_1}, F(F(\boldsymbol{X})) \in H_{s_2}, \boldsymbol{X} = F^{-1}(F(\boldsymbol{X})) \in H_{s_0}$, 等等. 这就说明 $F(\boldsymbol{X})$ 的旅程就是 $(\cdots s_{-1}s_0 \cdot s_1 s_2 \cdots)$, 因而

$$S(F(\boldsymbol{X})) = (\cdots s_{-1}s_0 \cdot s_1 s_2 \cdots) = \sigma(S(\boldsymbol{X})),$$

这就是我们要的共轭方程. 我们将 S 为同胚的证明留给读者 (见习题 3).

上一章中对单边移位成立的所有性质对现在的双边移位也都成立. 例如, σ 正好有 2^n 个周期为 n 的周期点, σ 也有一条稠密轨道, 等等. 而且 F 在 Λ 上是混沌的 (见习题 4 和习题 5). 但是, 现在还出现了一些新现象. 我们称两个点 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 是**前向渐近的**, 如果对所有的 $n \geqslant 0$, 都有 $F^n(\boldsymbol{X}_1), F^n(\boldsymbol{X}_2) \in D$, 并且

$$\lim_{n\to\infty} \left|F^n(\boldsymbol{X}_1) - F^n(\boldsymbol{X}_2)\right| = 0.$$

同样地, 两个点 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 是**后向渐近的**, 如果它们的后向轨道对所有的 n 都有定义, 并且当 $n \to -\infty$ 时, 上面的极限为零. 直观上, 如果两个点的前向轨道在 $n \to \infty$ 时彼此接近, 则这两个点是前向渐近的. 注意, 所有在 F 前向迭代下离开 S 的点都是与不动点 $\boldsymbol{X}_0 \in D_1$ 前向渐近的. 而且, 如果 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 位于 Λ_+ 中的同一条水平线段上, 则它们是前向渐近的; 如果 $\boldsymbol{X}_1$ 和 $\boldsymbol{X}_2$ 位于 Λ_- 中的同一条竖直线段上, 则它们是后向渐近的.

我们定义 $\boldsymbol{X}$ 的**稳定集** 为

$$W^s(\boldsymbol{X}) = \left\{\boldsymbol{Z} \middle| \text{当 } n \to \infty \text{ 时}, \ |F^n(\boldsymbol{Z}) - F^n(\boldsymbol{X})| \to 0\right\}.$$

类似地, $\boldsymbol{X}$ 的**不稳定集**定义为

$$W^u(\boldsymbol{X}) = \left\{\boldsymbol{Z} \middle| \text{当 } n \to \infty \text{ 时}, \ |F^{-n}(\boldsymbol{Z}) - F^{-n}(\boldsymbol{X})| \to 0\right\},$$

换句话说就是, 一个点 $\boldsymbol{Z}$ 位于 $W^s(\boldsymbol{X})$ 中当且仅当 $\boldsymbol{Z}$ 和 $\boldsymbol{X}$ 是前向渐近的. 与上面一样, 在马蹄映射的前向迭代下离开 S 的点都位于不动点 $\boldsymbol{X}_0 \in D_1$ 的稳定集上.

Λ 中点的稳定集和不稳定集则要复杂得多. 例如, 考虑位于 H_0 中的不动点 $\boldsymbol{X}^*$, 显然它的旅程为 $(\cdots 00 \cdot 000 \cdots)$. 经过 $\boldsymbol{X}^*$ 的水平线段 l_s 上的任一点都在 $W^s(\boldsymbol{X}^*)$ 中. 但是在这个稳定集中还有许多其他的点. 假设点 $\boldsymbol{Y}$ 最终要被映入 l_s 中, 于是存在自然数 n 使得 $|F^n(\boldsymbol{Y}) - \boldsymbol{X}^*| < 1$, 从而

$$|F^{n+k}(\boldsymbol{Y}) - \boldsymbol{X}^*| < \delta^k,$$

由此可得 $\boldsymbol{Y} \in W^s(\boldsymbol{X}^*)$. 因而所有由 $F^{-k}(l_s), k = 1, 2, 3, \cdots$ 给出的水平线段的并都位于 $W^s(\boldsymbol{X}^*)$ 中. 读者可以容易地验证有 2^k 条这样的水平线段.

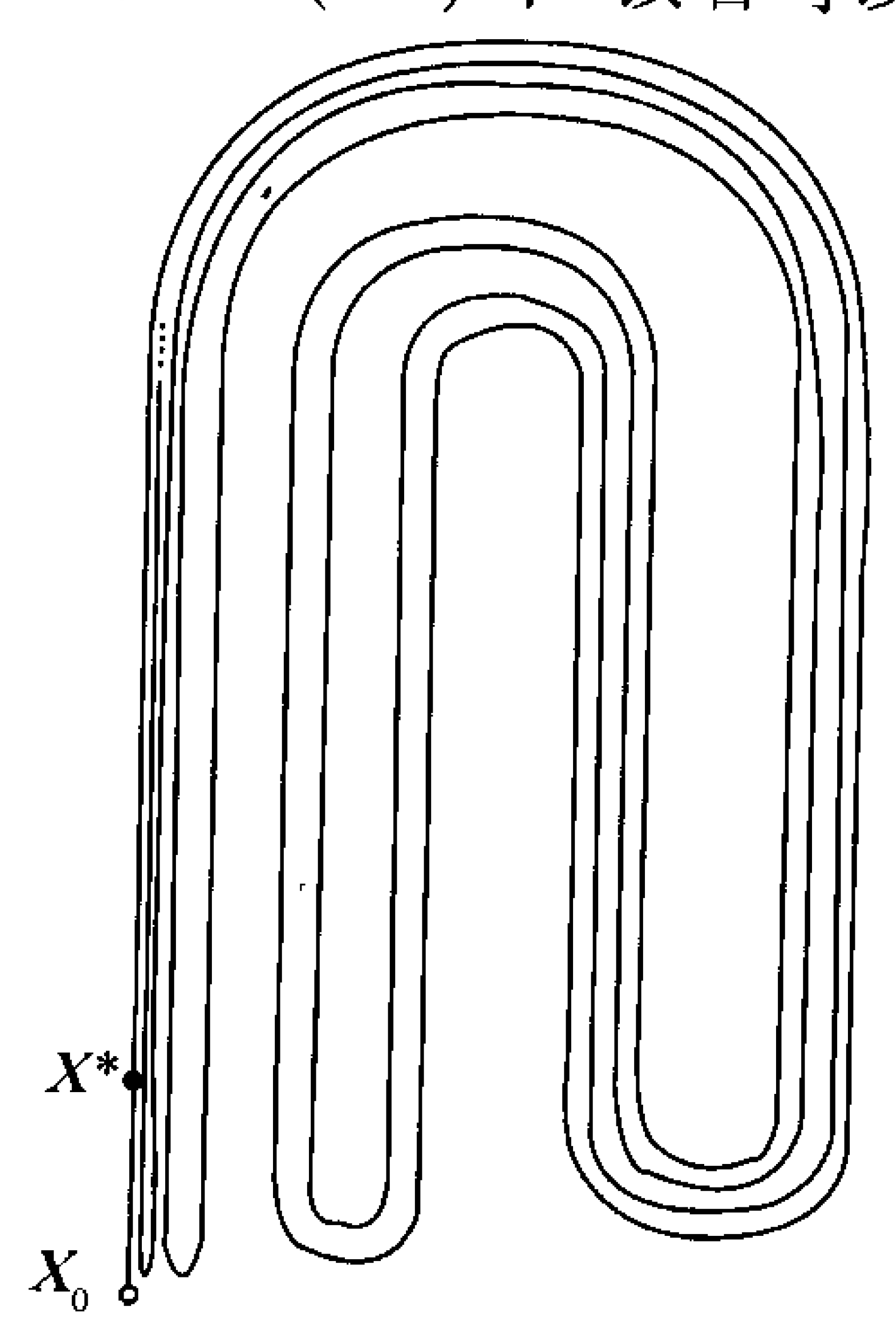


图 16.9 $\boldsymbol{X}^*$ 在 D 中的不稳定集

由于 $F(D) \subset D$, 不动点 $\boldsymbol{X}^*$ 的不稳定集的形式则稍有不同. D 中过 $\boldsymbol{X}^*$ 的竖直线段 l_u 显然位于 $W^u(\boldsymbol{X}^*)$ 中. 与上面一样, 再加上有 $F(D) \subset D$, l_u 在所有前向迭代下的像都位于 D 中. 容易验证 $F^k(l_u)$ 是 D 中的一条 "蛇形" 曲线, 它沿竖直方向贯穿 S 正好 2^k 次 (见图 16.9). 因而这些前向迭代像的并是一条非常复杂的曲线, 它要无穷次地贯穿 S. 这条曲线的闭包事实上包含 Λ 中的所有点以及它们的不稳定曲线 (见习题 12).

用移位映射的语言可以很容易地描述这些稳定集和不稳定集位于 Λ 中的部分. 令

$$\mathbf{s}^* = (\cdots s_{-2}^* s_{-1}^* \cdot s_0^* s_1^* s_2^* \cdots) \in \Sigma_2.$$

显然, 如果在某个位置的所有右边的位置上, 序列 $\mathbf{t}$ 的元素都与 $\mathbf{s}^*$ 的元素相同, 则 $\mathbf{t} \in W^s(\mathbf{s}^*)$. 这一陈述的逆命题也成立, 其证明参见习题 6.

现在, 一个自然的问题就是：一维合理映射中的集合 Λ 与马蹄映射中的集合 Λ 到底是一种什么关系? 直觉上, 似乎马蹄中的 Λ 包含更多的点. 但事实上, 这两个 Λ 是同胚的! 从移位映射可以清楚地看出这一点.

记 Σ_2^1 为所有单边 0, 1 序列构成的集合, 仍然用 Σ_2 记所有双边 0, 1 序列构成的集合. 定义一个映射

$$\Phi : \Sigma_2^1 \to \Sigma_2,$$

它的具体形式为

$$\Phi(s_0 s_1 s_2 \cdots) = (\cdots s_5 s_3 s_1 \cdot s_0 s_2 s_4 \cdots).$$

容易验证, Φ 是 Σ_2^1 和 Σ_2 之间的一个同胚 (见习题 11).

最后, 我们回头看看 16.1 节中的对象. 注意, 上节中研究的返回映射涉及无穷多片, 而在每一片上, 返回映射都类似于这一节的马蹄映射. 当然, 在马蹄映射发生混沌的区域上, 映射本质上是线性的, 因而本节的结果不能立刻就用来证明上节中同宿轨附近的返回映射也具有类似的性质. 但这是可以做到的, 只不过这样做所需要的技巧 (涉及一种推广的双曲性概念) 超出了本书的范围. 其中的具体细节请参见 [13] 或 [36].

16.3 双螺线吸引子

本节我们继续研究三维系统中同宿解附近的行为. 我们还是回到 16.1 节中描述的系统, 不过现在假设向量场是关于原点反称的. 特别地, 这将意味着原点处不稳定曲线的两支 $\zeta^{\pm}$ 都是如图 16.10 所示的同宿解. 我们仍然假设 ζ^{+} 与圆柱面 $C: r=1, |z| \leqslant 1$ 交于点 $\theta=0, z=0$, 因而 ζ^{-} 就与圆柱面交于对径点 $\theta=\pi, z=0$.

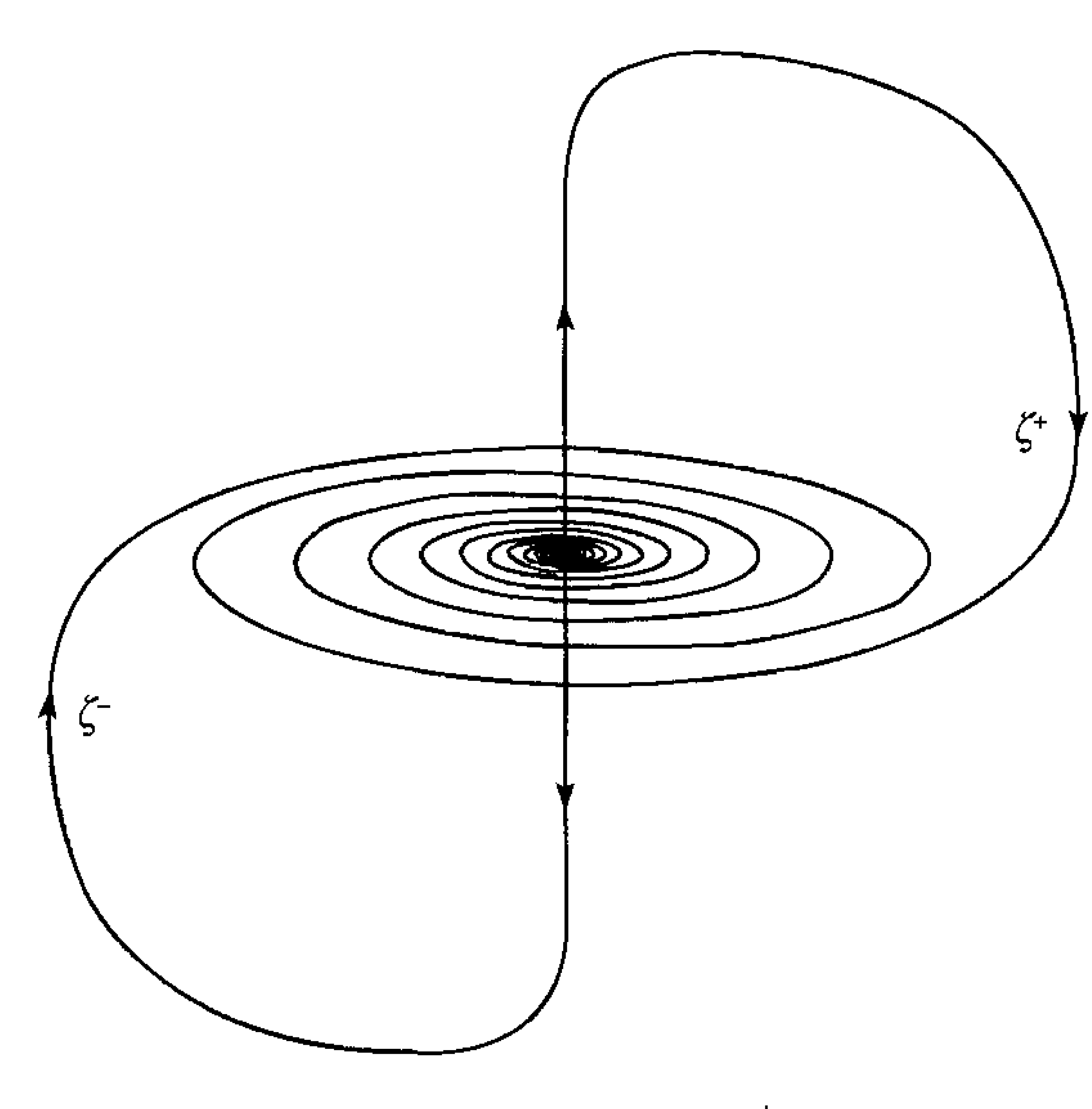


图 16.10 同宿轨 $\zeta^{\pm}$

同 16.1 节中一样, 我们可以得到一个定义在圆柱面 C 上的庞加莱映射 Φ. 不过, 这次我们不能再忽略到达 C 中 $z<0$ 的区域中的解; 现在这些解将跟随第二个同宿解 ζ^{-} 前进, 直到与 C 再次相交. 因而 Φ 在整个 $C-\{z=0\}$ 上都有定义.

与前面一样, 定义在圆柱面上半部分 C^{+} 的庞加莱映射 Φ^{+} 为

$$\Phi^{+}\begin{pmatrix}\theta_0\\ z_0\end{pmatrix}=\begin{pmatrix}\theta_1\\ z_1\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\sqrt{z_0}\sin\left(\theta_0+\ln\sqrt{z_0}\right)\\ \frac{1}{2}\sqrt{z_0}\cos\left(\theta_0+\ln\sqrt{z_0}\right)\end{pmatrix}.$$

利用向量场的对称性, 可以算出定义在 C^{-} 上的 Φ^{-} 为

$$\Phi^{-}\begin{pmatrix}\theta_0\\ z_0\end{pmatrix}=\begin{pmatrix}\theta_1\\ z_1\end{pmatrix}=\begin{pmatrix}\pi-\frac{1}{2}\sqrt{-z_0}\sin\left(\theta_0+\ln\sqrt{-z_0}\right)\\ \frac{1}{2}\sqrt{-z_0}\cos\left(\theta_0+\ln\sqrt{-z_0}\right)\end{pmatrix},$$

其中 $z_0 < 0$, 而 θ_0 则是任意的. 于是, $\Phi(C^+)$ 就是中心在 $\theta = 0, z = 0$, 半径为 1/2 的圆盘, 而 $\Phi(C^-)$ 就是中心在 $\theta = \pi, z = 0$ 的一个类似圆盘. 由于 $\zeta^\pm$ 正好就从这两个圆盘的中心进入 C, 因此, 中心并不在 Φ 的像中 (见图 16.11).

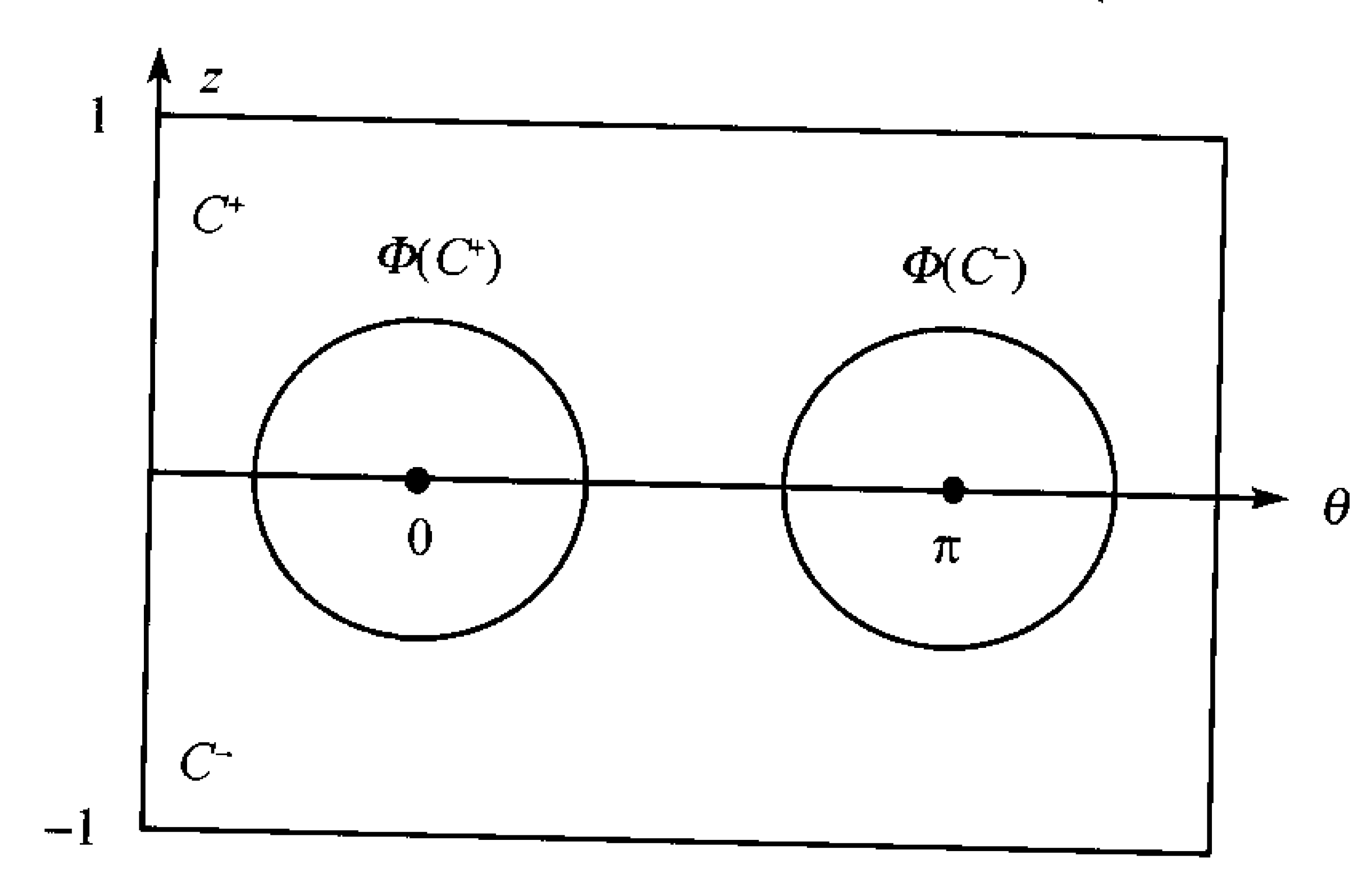


图 16.11 $\Phi(C^\pm) \cap C$, 这里我们将圆柱面 C 画成了长条

现在令 $\boldsymbol{X} \in C$, 则要么过 $\boldsymbol{X}$ 的解位于原点的稳定曲面上, 要么 $\Phi(\boldsymbol{X})$ 有定义, 因而过 $\boldsymbol{X}$ 的解将在后面某个时间返回到 C. 作为一个推论, C 中的每个点 $\boldsymbol{X}$ 都具有如下的性质:

(1) 要么在 $t \to \infty$ 时, 过 $\boldsymbol{X}$ 的解无穷次地穿过 C, 从而对所有的 $n \geqslant 0$, $\Phi^n(\boldsymbol{X})$ 都有定义;

(2) 要么过 $\boldsymbol{X}$ 的解最终与 $z = 0$ 相交, 因而它位于过原点的稳定曲面上.

当时间向后时, 情形则不同: 只有位于 $\Phi(C^\pm)$ 上的那些点才会返回到 C; 严格地说, 我们并没有定义 $C - \Phi(C^\pm)$ 中点的向后解, 但是我们认为这些解在 $\mathbb{R}^3$ 中都有定义并且最终会与 C 相交, 而且在此之后会连续地与 C 相交.

与洛伦茨吸引子情形一样, 令

$$A = \bigcap_{n=0}^{\infty} \overline{\Phi^n(C)},$$

其中 $\overline{\Phi^n(C)}$ 代表集合 $\Phi^n(C)$ 的闭包. 然后, 令

$$\mathcal{A} = \left(\bigcup_{t \in \mathbb{R}} \phi_t(A) \right) \bigcup \{(0,0,0)\}.$$

注意, $\overline{\Phi^n(C)} - \Phi^n(C)$ 正好就是同宿解 $\zeta^\pm$ 与 C 的两个交点. 因而为了保证 $\mathcal{A}$ 是一个闭集, 只需要加上原点就可以了.

下一结果的证明思想与 14.4 节证明洛伦茨吸引子的相应结果是类似的.

命题 集合 $\mathcal{A}$ 具有下面的性质:

(1) $\mathcal{A}$ 是紧不变的;

(2) 存在一个包含 $\mathcal{A}$ 的开集 U 使得对每个 $\boldsymbol{X} \in U$ 以及所有的 $t \geqslant 0$, 都有 $\phi_t(\boldsymbol{X}) \in U$, 并且 $\cap_{t\geqslant 0}\phi_t(U) = \mathcal{A}$. ■

因而除了传递性外, $\mathcal{A}$ 具有一个吸引子的所有性质. 尽管如此, 习惯上仍然称 $\mathcal{A}$ 为一个双螺线吸引子.

由于没有写出这个系统的表达式, 我们不能像洛伦茨情形那样去计算双螺线向量场的散度. 但是, 我们有庞加莱映射 Φ 的表达式. 直接的计算表明 $\det D\Phi = 1/8$. 即, 每次迭代后, 庞加莱映射 Φ 将面积缩小到原来的 1/8. 于是, $A = \cap_{n\geqslant 0}\overline{\Phi^n(C)}$ 在 C 中的面积为 0, 由此可得:

命题 双螺线吸引子 $\mathcal{A}$ 的体积为零. ■

16.4 同宿分岔

在高维, 与同宿轨有关的分岔导致的行为之复杂令人恐怖 (或者说美妙, 在于你怎么去看). 在本节, 我们简要地介绍几种这类分岔. 这里我们将讨论双螺线向量场一个特定的扰动, 这个扰动将打断两个同宿连接.

这个分岔的完整图像涉及如何理解无穷多个马蹄映射的"开折". 这句话的含义如下. 考虑定义在矩形 R 上的一族映射 F_λ, 其中参数 $\lambda \in [0,1]$. 像集 $F_\lambda(R)$ 是一个如图 16.12 所示的马蹄. 当 $\lambda = 0$ 时, $F_\lambda(R)$ 位于 R 的下方. 当 λ 增加时, $F_\lambda(R)$ 单调地上升. 当 $\lambda = 1$ 时, $F_\lambda(R)$ 穿过 R 两次, 并且我们假设 F_1 就是 16.2 节中所描述的马蹄映射.

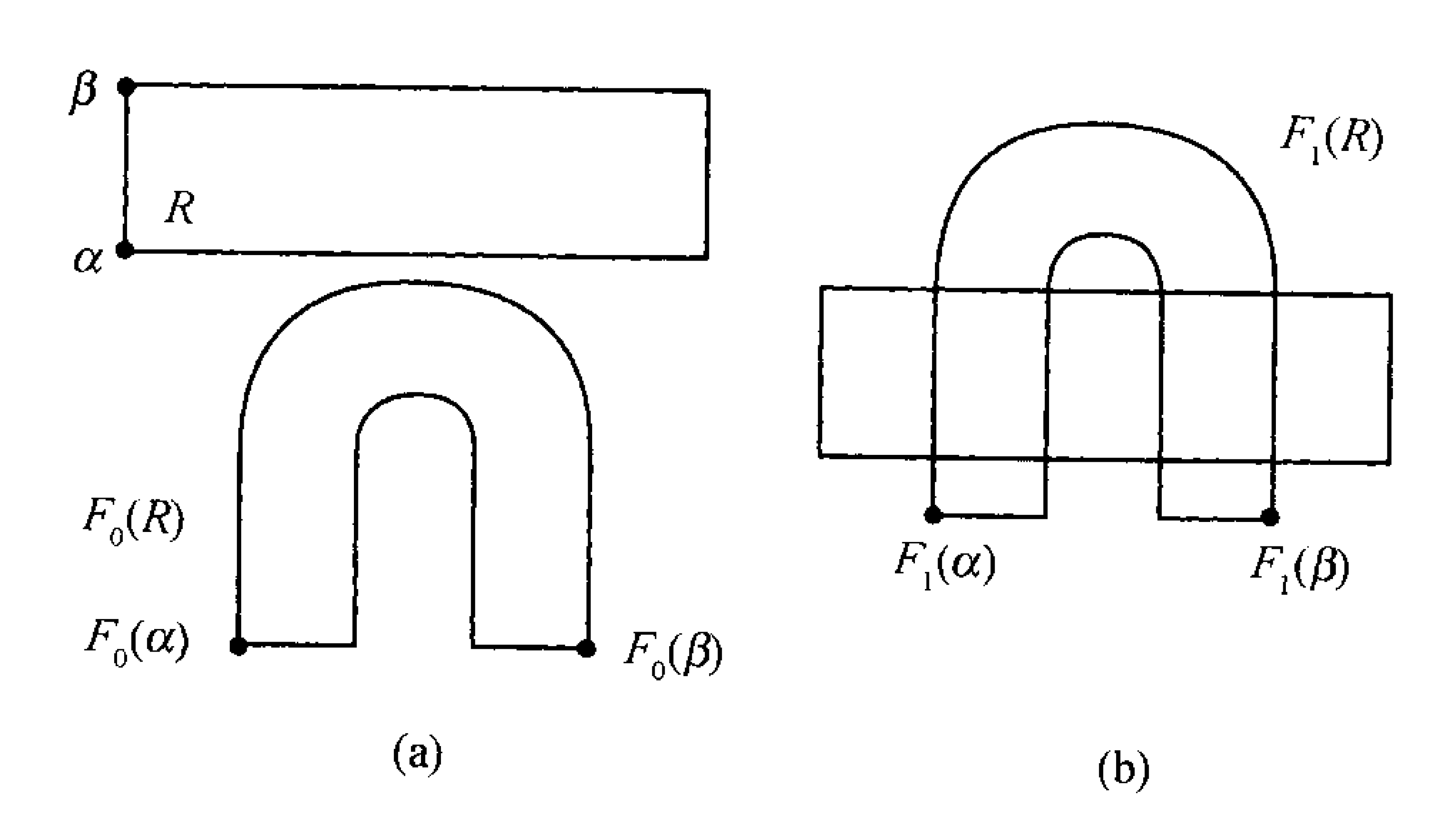


图 16.12 在 (a) $\lambda = 0$ 和 (b) $\lambda = 1$ 时的像集 $F_\lambda(R)$

显然, F_0 在 R 中没有任何周期点, 但是当 λ 到达 1 时, 却产生无穷多个周期点, 而且还将出现其他的混沌行为. 在通往马蹄映射的途中, 映射族 F_λ 将经历无穷多次分岔. 这些分岔如何出现是数学中非常现代的研究课题.

相对于一维合理映射族 $f_\lambda(x) = \lambda x(1-x)$ $(0 \leqslant \lambda \leqslant 4)$ 中出现的分岔, 现在的

情形要复杂得多. 合理映射族的分岔结构在最近已经被完全确定; 而解决现在这种平面情形依然遥遥无期.

现在我们在双螺线系统中引入一个参数 ϵ. 当 $\epsilon = 0$ 时, 系统就是上节中考虑的双螺线系统. 当 $\epsilon \neq 0$ 时, 我们简单地将 $\zeta^+ \cap C$ (以及相应的转移映射) 沿 z 方向移动 ϵ 从而改变系统. 说得准确些, 我们假设系统在圆柱形区域 $r \leqslant 1, |z| \leqslant 1$ 中保持不变, 但是我们将定义在上半圆盘 D^+ 的转移映射的像加上 $(0, \epsilon)$, 即, 定义在 C^+ 上的新庞加莱映射为

$$\Phi_\epsilon^+ \begin{pmatrix} \theta \\ z \end{pmatrix} = \begin{pmatrix} \dfrac{1}{2}\sqrt{z}\sin\left(\theta + \ln\sqrt{z}\right) \\ \dfrac{1}{2}\sqrt{z}\cos\left(\theta + \ln\sqrt{z}\right) + \epsilon \end{pmatrix},$$

并且利用系统的反称性类似地定义 Φ_ϵ^-. 而且我们还假设 ϵ 充分小 ($|\epsilon| < 1/2$) 使得 $\Phi_\epsilon^\pm(C) \subset C$.

当 $\epsilon > 0$ 时, ζ^+ 与 C 交于上方的圆柱面 C^+, 然后在经过原点附近后, 它将缠绕着自己前进直到与 C 第二次再相交. 当 $\epsilon < 0$ 时, ζ^+ 与 C 交于 C^-, 而且它将沿一种非常不同的路径回到 C: 这次它缠绕 ζ^-.

注意, $\Phi_0^\pm$ 在 $C^\pm$ 中有无穷多个不动点. 当 $\epsilon \neq 0$ 时, 情况会非常不同.

命题　当 $\epsilon \neq 0$ 时, 每个映射 $\Phi_\epsilon^\pm$ 在 $C^\pm$ 中都只有有限个不动点.

证明　为了找出 Φ_ϵ^+ 的不动点, 需要求解

$$\theta = \frac{1}{2}\sqrt{z}\sin\left(\theta + \ln\sqrt{z}\right),$$
$$z = \frac{1}{2}\sqrt{z}\cos\left(\theta + \ln\sqrt{z}\right) + \epsilon,$$

其中 $\epsilon > 0$. 与 16.1 节一样, 必定有

$$\frac{z}{4} = \theta^2 + (z - \epsilon)^2,$$

从而

$$\theta = \pm\frac{1}{2}\sqrt{z - 4(z - \epsilon)^2}.$$

特别地, 必定有

$$z - 4(z - \epsilon)^2 \geqslant 0,$$

也就是说

$$\frac{4(z - \epsilon)^2}{z} \leqslant 1.$$

只有当 z 位于区间 I_ϵ 时, 上面的不等式才成立, 其中区间 I_ϵ 如下给出:

$$\frac{1}{8}+\epsilon-\frac{1}{8}\sqrt{1+16\epsilon}\leqslant z\leqslant\frac{1}{8}+\epsilon+\frac{1}{8}\sqrt{1+16\epsilon}.$$

为了保证 Φ_ϵ^+ 有不动点, 这就对 ϵ 加上了另外一个限制, 即 $\epsilon>-1/16$. 注意, 当 $\epsilon>-1/16$ 时, 有

$$\frac{1}{8}+\epsilon-\frac{1}{8}\sqrt{1+16\epsilon}>0,$$

从而 I_ϵ 的长度为 $\sqrt{1+16\epsilon}/4$, 并且这个区间在 0 的右边.

为了确定不动点的 z 值, 现在需要求解

$$\cos\left(\pm\frac{1}{2}\sqrt{z-4(z-\epsilon)^2}+\ln\sqrt{z}\right)=\frac{2(z-\epsilon)}{\sqrt{z}},$$

或者

$$\cos^2\left(\pm\frac{1}{2}\sqrt{z-4(z-\epsilon)^2}+\ln\sqrt{z}\right)=\frac{4(z-\epsilon)^2}{z}.$$

经过一点计算后可以验证函数

$$g(z)=\frac{4(z-\epsilon)^2}{z}$$

在 $z=\epsilon$ 处有唯一的极小值 0, 以及在区间 I_ϵ 端点处的两个极大值 1. 同时, 函数

$$h(z)=\cos^2\left(\pm\frac{1}{2}\sqrt{z-4(z-\epsilon)^2}+\ln\sqrt{z}\right)$$

的图像在区间 I_ϵ 中在 ±1 之间只振动有限次. 因而在 I_ϵ 中只有有限个 z 值满足 $h(z)=g(z)$. 这有限个点就是 Φ_ϵ^+ 的不动点. ■

注意, 当 $\epsilon\to0$ 时, 区间 I_ϵ 趋于 $[0,1/4]$, 从而 h 在 I_ϵ 中振动的次数将无限增加. 由此可得

推论 给定 $N\in\mathbb{Z}$, 存在 $\epsilon_N>0$ 使得只要 $0<\epsilon<\epsilon_N$, Φ_ϵ^+ 在 C^+ 中就至少有 N 个不动点. ■

当 $\epsilon>0$ 时, 不稳定曲线 ζ^+ 在第一次穿过 C 时与稳定曲面并不相交. 事实上, 此时 ζ^+ 与 C^+ 交于 $\theta=0, z=\epsilon$. 但这并不意味着在 $\epsilon\neq0$ 时就没有同宿轨. 事实上, 我们有下面的命题:

命题 存在无穷多个 ϵ 值, 使得 $\zeta^\pm$ 是两次穿过 C 的同宿解.

证明 为了证明这一点, 需要找出一些 ϵ 的取值, 使得 $\Phi_\epsilon^+(0,\epsilon)$ 位于原点的稳定曲面上. 因而需要求解

$$0=\frac{\sqrt{\epsilon}}{2}\cos\left(0-\ln\sqrt{\epsilon}\right)+\epsilon,$$

即

$$-2\sqrt{\epsilon} = \cos\big(-\ln\sqrt{\epsilon}\big).$$

但是, 与 16.1 节一样, $\cos(-\ln\sqrt{\epsilon})$ 的图像与 $-2\sqrt{\epsilon}$ 的图像相交无穷多次. ■

对于每个使得 ζ^+ 为同宿解的 ϵ 值, 又有无穷多个不动点 (对 $\Phi_\epsilon^\pm \circ \Phi_\epsilon^\pm$), 同时得到一个结构相当不同的吸引子. 显然, 当 ϵ 变化时, 发生了很多事情. 我们邀请学完本书的读者继续努力, 理解这里发生的每件事. 祝你好运! 玩得开心!

16.5　探索: Chua 电路

在这个探索中, 我们来研究一个非线性的三维微分方程系统, 这个系统与电子电路中的 Chua 电路有关. 这是第一个表现出混沌行为的电路方程组的例子. 事实上, 当参数取某些值时, 这个方程组表现出的行为类似于 16.3 节中的双螺线吸引子. 原始的 Chua 电路方程的非线性来自于分片线性. 这里我们研究这个方程组的一个变种, 它的非线性由一个三次函数给出. 关于 Chua 电路更加详细的细节, 参见 [11] 和 [25].

非线性的 Chua 电路系统为

$$\begin{aligned} x' &= a(y-\phi(x)), \\ y' &= x - y + z, \\ z' &= -by, \end{aligned}$$

其中 a, b 为参数, 而函数 ϕ 的表达式为

$$\phi(x) = \frac{1}{16}x^3 - \frac{1}{6}x.$$

事实上, 这个多项式的系数通常是以参数形式出现的, 但是为了确定起见, 在这个探索中, 我们将它们固定. 当 $a = 10.918\,65\cdots$ 且 $b = 14$ 时, 这个系统似乎有两条对称的同宿轨, 如图 16.13 所示. 这个探索的目的就是研究当参数 a 变化时, 系统将如何变化. 因而, 我们将参数 b 固定在 14, 而让参数 a 变化. 我们提醒做实验的读者注意, 证明下面观察到的任何混沌行为或分岔现象都是几乎不可能的; 事实上, 你在这方面所做的任何事情都将是一个合格且有趣的研究工作.

(1) 同前面一样, 先找出所有的平衡点.

(2) 确定这些平衡点的类型, 此时也许需要借助一种计算机代数系统.

(3) 该系统具有一种对称性; 描述这种对称性并指出这对系统的解蕴涵什么.

(4) 让 a 从 6 变化到 14. 描述当 a 变化时你观察到的一切分岔. 为了看出分岔的全局效应, 请务必在这个实验中 (包括后面的其他实验) 选取一对位置对称的初值条件. 请特别关注从原点附近出发的解.

(5) 是否存在一些 a 值使得系统看起来有一个吸引子? 此时看起来发生了什么? 你能否构造一个模型?

(6) 描述下列 a 值附近出现的分岔:

(a) $a = 6.58$; (b) $a = 7.3$; (c) $a = 8.78$; (d) $a = 10.77$.

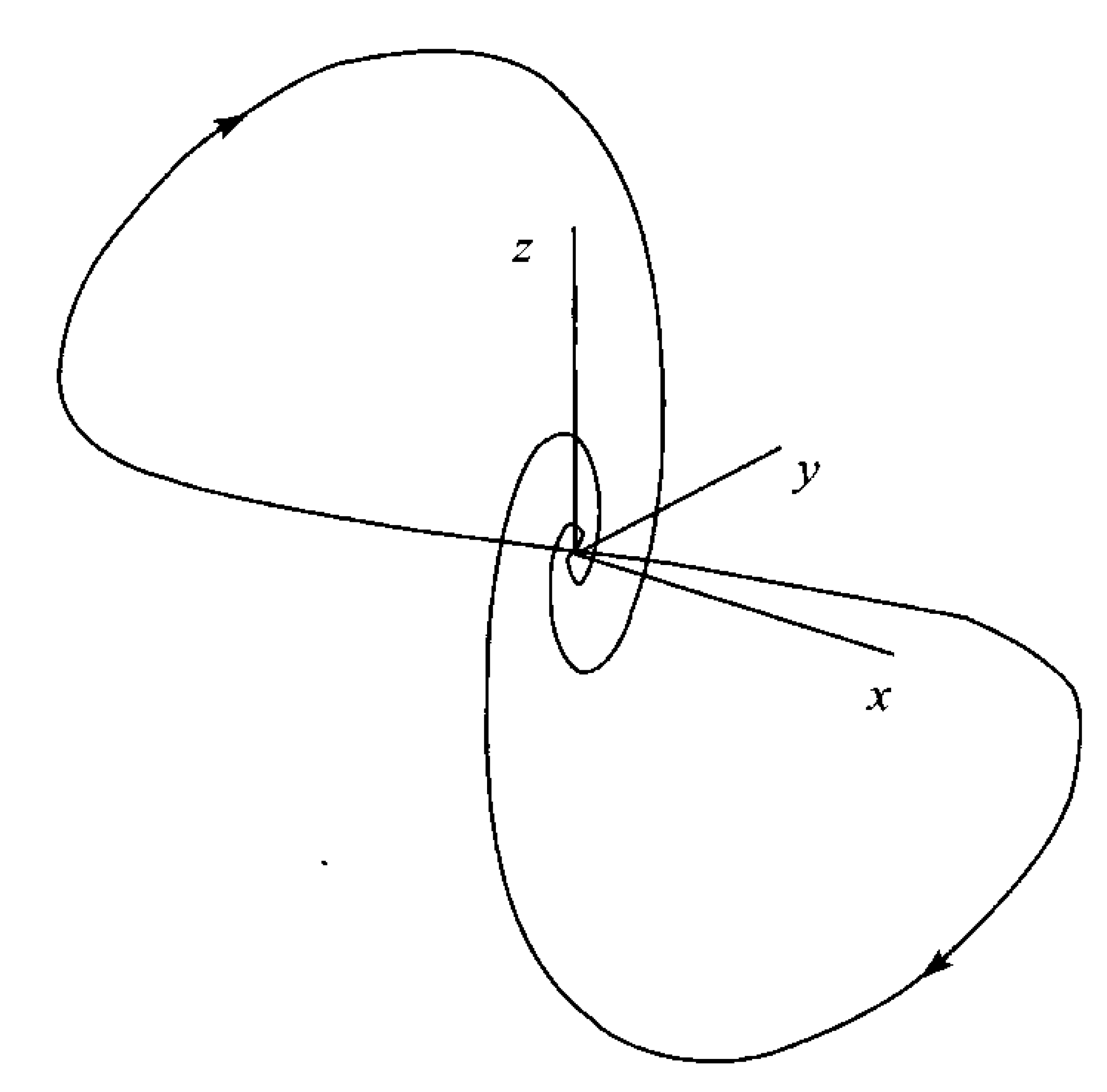


图 16.13　在参数值 $a = 10.918\ 65\cdots$ 且 $b = 14$ 时非线性 Chua 电路中的两条同宿轨

习　题

1. 证明
$$d((\mathbf{s}),(\mathbf{t})) = \sum_{i=-\infty}^{\infty} \frac{|s_i - t_i|}{2^{|i|}}$$
是 Σ_2 上的一个距离函数, 其中 Σ_2 就是 16.2 节中描述的双边无穷 0, 1 序列构成的集合.
2. 证明移位映射 σ 是 Σ_2 上的一个同胚.
3. 证明 $S : \Lambda \to \Sigma_2$ 给出了 σ 和 F 之间的一个共轭.
4. 构造 σ 的一条稠密轨道.
5. 证明 σ 的周期点稠密.
6. 令 $\mathbf{s}^* \in \Sigma_2$. 证明 $W^s(\mathbf{s}^*)$ 正好由那些某个位置的所有右边的位置上的元素都与 $\mathbf{s}^*$ 的元素相同的序列构成.
7. 令 $(\mathbf{0}) = (\cdots 00 \cdot 000 \cdots) \in \Sigma_2$. 序列 $\mathbf{s} \in \Sigma_2$ 称为**同宿**于 $(\mathbf{0})$, 如果 $\mathbf{s} \in W^s(\mathbf{0}) \cap W^u(\mathbf{0})$. 描述一个同宿于 $(\mathbf{0})$ 的序列的元素. 证明同宿于 $(\mathbf{0})$ 的序列在 Σ_2 中稠密.
8. 令 $(\mathbf{1}) = (\cdots 11 \cdot 111 \cdots) \in \Sigma_2$. 如果 $\mathbf{s} \in W^s(\mathbf{0}) \cap W^u(\mathbf{1})$, 则称序列 $\mathbf{s}$ 是一个异

宿序列. 描述一个这样异宿序列的元素. 证明所有这样的异宿序列在 Σ_2 中稠密.

9. 将同宿点和异宿点的定义推广到 σ 的任何周期点. 并且在现在的情形重新证明习题 7 和习题 8 的结论.
10. 证明任一给定周期点的同宿点集合是可数的.
11. 记 Σ_2^1 为所有单边 0,1 序列构成的集合. 定义映射 $\Phi: \Sigma_2^1 \to \Sigma_2$ 为
$$\Phi(s_0 s_1 s_2 \cdots) = (\cdots s_5 s_3 s_1 \cdot s_0 s_2 s_4 \cdots).$$
证明 Φ 是一个同胚.
12. 记 $\boldsymbol{X}^*$ 为马蹄映射 F 在 H_0 中的不动点. 证明 $W^u(\boldsymbol{X}^*)$ 的闭包包含 Λ 的所有点以及它们的不稳定曲线上的所有点.
13. 令 $R: \Sigma_2 \to \Sigma_2$ 的定义为
$$R(\cdots s_{-2} s_{-1} \cdot s_0 s_1 s_2 \cdots) = (\cdots s_2 s_1 s_0 \cdot s_{-1} s_{-2} \cdots).$$
证明 $R \circ R = \mathrm{id}$ 和 $\sigma \circ R = R \circ \sigma^{-1}$. 并由此推出 $\sigma = U \circ R$, 其中 U 是一个满足 $U \circ U = \mathrm{id}$ 的映射. 如果一个映射本身就是它的逆, 则称之为一个**对合**. 它们代表一类非常简单的动力系统. 上面的结论表明移位映射可以分解成两个这种映射的复合.
14. 继续考虑上题中定义的映射 R, 令 s 是一个在 R 作用下不动的序列. 假设 $\sigma^n(\mathrm{s})$ 也在 R 的作用下不动. 证明 s 是 σ 的一个周期为 $2n$ 的周期点.
15. 现在假设 $\sigma^n(\mathrm{s})$ 在 U 的作用下不动. 重做上题, 其中 U 的定义见习题 13. s 的周期是多少?
16. 对于第 14 章中的洛伦茨系统, 用数值方法研究当 r 从 13.92 变化到 13.96 时出现的分岔, 此时 $\sigma = 10$, $b = 8/3$.

第 17 章　再论存在唯一性

本章我们回到第 7 章曾经出现的素材，这次我们将补充早先省略的所有技术细节和证明. 后果就是这一章比前面每一章都要困难; 但这是严格研究常微分方程的根本所在. 为了完全理解这里的许多证明, 读者必须熟悉实分析中的一致连续性、函数的一致收敛性以及紧集等概念.

17.1　存在唯一性定理

考虑自治微分方程系统

$$\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X}),$$

其中 $\boldsymbol{F}:\mathbb{R}^n \to \mathbb{R}^n$. 在以前的章节中, 我们通常假设 $\boldsymbol{F}$ 是 C^∞; 这里我们放宽这个条件, 只假设 $\boldsymbol{F}$ 是 C^1 的. 这意味着 $\boldsymbol{F}$ 是连续可微的. 即, 它的所有一阶偏导数都存在并且是 $\mathbb{R}^n$ 上的连续函数. 在本章的前几节, 我们将只讨论自治方程; 但之后我们将假设 $\boldsymbol{F}$ 还连续可微地依赖于 t 和 $\boldsymbol{X}$.

我们已经知道, 上述系统的一个解就是定义在某个区间 $J \subset \mathbb{R}$ 上的一个可微函数 $\boldsymbol{X}: J \to \mathbb{R}^n$, 使得对所有的 $t \in J$, 都有

$$\boldsymbol{X}'(t) = \boldsymbol{F}(\boldsymbol{X}(t)).$$

几何上, $\boldsymbol{X}(t)$ 是 $\mathbb{R}^n$ 中的一条曲线, 其切向量 $\boldsymbol{X}'(t)$ 等于 $\boldsymbol{F}(\boldsymbol{X}(t))$; 与前面各章一样, 我们认为这个向量的基点位于 $\boldsymbol{X}(t)$, 因而映射 $\boldsymbol{F}:\mathbb{R}^n \to \mathbb{R}^n$ 就定义了 $\mathbb{R}^n$ 上的一个向量场. 解 $\boldsymbol{X}: J \to \mathbb{R}^n$ 的一个**初始条件**或**初值**就是形如 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$ 的一个指定, 其中 $t_0 \in J, \boldsymbol{X}_0 \in \mathbb{R}^n$. 为了简单起见, 我们通常取 $t_0 = 0$.

一个非线性微分方程满足一个给定初始条件的解可能有好几个. 例如, 考虑一阶非线性微分方程

$$x' = 3x^{2/3}.$$

在第 7 章, 我们已经看到, 恒等于零的函数 $u_0:\mathbb{R} \to \mathbb{R}$, $u_0(t) \equiv 0$ 就是一个满足初始条件 $u(0) = 0$ 的解. 但是, $u_1(t) = t^3$ 也是满足上面初始条件的一个解, 此外对任一 $\tau > 0$, 函数

$$u_\tau(t) = \begin{cases} 0, & \text{如果 } t \leqslant \tau, \\ (t-\tau)^3, & \text{如果 } t > \tau, \end{cases}$$

也是满足初始条件 $u_\tau(0)=0$ 的一个解.

除了解的唯一性问题外, 还有解的存在性问题. 在讨论线性系统时, 我们能够将解显式地算出来. 对于非线性系统, 我们已经看到, 这常常是不可能的. 而且, 某些初始条件可能不会产生任何解. 例如, 我们在第 7 章已经看到, 微分方程

$$x' = \begin{cases} 1, & \text{如果 } x < 0, \\ -1, & \text{如果 } x \geqslant 0 \end{cases}$$

没有解满足 $x(0)=0$.

由此可见, 为了保证解的存在唯一性, 我们必须对函数 $\boldsymbol{F}$ 另外再加一些条件. 我们将看到, 假设 $\boldsymbol{F}$ 连续可微就足够了. 在上面的第一个例子中, $\boldsymbol{F}$ 在出问题的点 $x=0$ 处不可微, 而在第二个例子中, $\boldsymbol{F}$ 在 $x=0$ 处不连续.

下面就是常微分方程的基本局部定理.

存在唯一性定理 考虑初值问题

$$\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X}), \quad \boldsymbol{X}(0) = \boldsymbol{X}_0,$$

其中 $\boldsymbol{X}_0 \in \mathbb{R}^n$. 假设 $\boldsymbol{F}: \mathbb{R}^n \to \mathbb{R}^n$ 是 C^1 的. 则该初值问题存在唯一的解. 更准确地说, 存在 $a>0$, 以及该微分方程满足初值条件

$$\boldsymbol{X}(0) = \boldsymbol{X}_0$$

的唯一解

$$\boldsymbol{X}: (-a, a) \to \mathbb{R}^n.$$ ■

我们将在下节证明这个定理.

17.2 存在唯一性的证明

我们需要回忆一下多元微积分. 令 $\boldsymbol{F}: \mathbb{R}^n \to \mathbb{R}^n$. 如果 $\mathbb{R}^n$ 中的坐标为 $(x_1, \cdots, x_n)$, 则我们可以将 $\boldsymbol{F}(\boldsymbol{X})$ 写成

$$\boldsymbol{F}(\boldsymbol{X}) = (f_1(x_1, \cdots, x_n), \cdots, f_n(x_1, \cdots, x_n)).$$

记 $D\boldsymbol{F}_{\boldsymbol{X}}$ 为 $\boldsymbol{F}$ 在 $\boldsymbol{X} \in \mathbb{R}^n$ 处的导算子. 我们可以从两种稍微不同的观点来看导算子 $D\boldsymbol{F}_{\boldsymbol{X}}$. 一种观点是, 对每个 $\boldsymbol{X} \in \mathbb{R}^n$, $D\boldsymbol{F}_{\boldsymbol{X}}$ 都是一个线性映射; 这个线性映射将 $\mathbb{R}^n$ 中的一个向量 $\boldsymbol{U}$ 映到向量

$$D\boldsymbol{F}_{\boldsymbol{X}}(\boldsymbol{U}) = \lim_{h\to 0} \frac{\boldsymbol{F}(\boldsymbol{X}+h\boldsymbol{U}) - \boldsymbol{F}(\boldsymbol{X})}{h},$$

其中 h 在 $\mathbb{R}$ 中变化. 等价地, 从矩阵观点来看, DF_X 就是 $n \times n$ 的雅可比矩阵

$$DF_X = \left(\frac{\partial f_i}{\partial x_j}\right),$$

其中所有的偏导数都在 $(x_1, \cdots, x_n)$ 处取值. 因而, 导算子还可以看成是一个映射, 它将 $\mathbb{R}^n$ 中的每个点映到一个线性映射或矩阵, 即 $DF : \mathbb{R}^n \to L(\mathbb{R}^n)$.

与前面一样, 函数 F 称为连续可微的, 或 C^1 的, 如果这些 f_j 的所有偏导数都存在并且连续. 我们假设, 在本章余下的部分, F 都是 C^1 的. 对每个 $X \in \mathbb{R}^n$, 我们定义雅可比矩阵 DF_X 的范数 $|DF_X|$ 为

$$|DF_X| = \sup_{|U|=1} |DF_X(U)|,$$

其中 $U \in \mathbb{R}^n$. 注意, $|DF_X|$ 不一定就是 X 点处的雅可比矩阵的最大特征值 (在绝对值意义下).

例 假设

$$DF_X = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

则事实上有 $|DF_X| = 2$, 而且 2 就是 DF_X 的最大特征值. 然而, 如果

$$DF_X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

则

$$\begin{aligned} |DF_X| &= \sup_{0 \leqslant \theta \leqslant 2\pi} \left| \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \right| \\ &= \sup_{0 \leqslant \theta \leqslant 2\pi} \sqrt{(\cos\theta + \sin\theta)^2 + \sin^2\theta} \\ &= \sup_{0 \leqslant \theta \leqslant 2\pi} \sqrt{1 + 2\cos\theta\sin\theta + \sin^2\theta} \\ &> 1, \end{aligned}$$

而这里 1 是最大的特征值. ■

但是, 对 $\mathbb{R}^n$ 中的任一向量 V, 我们都有

$$|DF_X(V)| \leqslant |DF_X|\ |V|.$$

事实上, 如果将 V 写成 $V = (V/|V|)|V|$, 则由于 $V/|V|$ 的长度为 1, 我们就有

$$|DF_X(V)| = \big|DF_X\big(V/|V|\big)\big||V| \leqslant |DF_X|\ |V|.$$

而且, $\boldsymbol{F}:\mathbb{R}^n\to\mathbb{R}^n$ 为 C^1 蕴涵将 $\boldsymbol{X}$ 映到 $D\boldsymbol{F}_{\boldsymbol{X}}$ 的映射 $D\boldsymbol{F}:\mathbb{R}^n\to L(\mathbb{R}^n)$ 是一个连续映射.

设 $\mathcal{O}\subset\mathbb{R}^n$ 是一个开集. 函数 $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 称为在 $\mathcal{O}$ 上是**利普希茨**的, 如果存在常数 $K>0$ 使得对所有的 $\boldsymbol{X},\boldsymbol{Y}\in\mathcal{O}$, 都有

$$|\boldsymbol{F}(\boldsymbol{Y})-\boldsymbol{F}(\boldsymbol{X})|\leqslant K|\boldsymbol{Y}-\boldsymbol{X}|.$$

我们称 K 是 $\boldsymbol{F}$ 的一个**利普希茨常数**. 更一般地, 我们称 $\boldsymbol{F}$ 是**局部利普希茨**的, 如果 $\mathcal{O}$ 中的每个点都有一个在 $\mathcal{O}$ 中的邻域 $\mathcal{O}'$, 使得 $\boldsymbol{F}$ 限制在 $\mathcal{O}'$ 上是利普希茨的. $\boldsymbol{F}|\mathcal{O}'$ 的利普希茨常数可能会随着 $\mathcal{O}'$ 的不同而变化.

另外一个重要的概念就是紧性. 我们称一个集合 $\mathcal{C}\subset\mathbb{R}^n$ 是**紧**的, 如果 $\mathcal{C}$ 是闭的, 并且有界. 一个重要的事实是, 如果 $f:\mathcal{C}\to\mathbb{R}$ 连续, 而且 $\mathcal{C}$ 紧, 则首先 f 在 $\mathcal{C}$ 上有界, 其次 f 在 $\mathcal{C}$ 上可以达到最大值. 见本章的习题 13.

引理　如果函数 $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 是 C^1 的, 则 $\boldsymbol{F}$ 是局部利普希茨的.

证明　假设函数 $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 是 C^1 的, 任取一点 $\boldsymbol{X}_0\in\mathcal{O}$, 再取 $\epsilon>0$ 充分小, 使得 $\boldsymbol{X}_0$ 的半径为 ϵ 的闭球 $\mathcal{O}_\epsilon$ 包含在 $\mathcal{O}$ 中. 令 K 是 $|D\boldsymbol{F}_{\boldsymbol{X}}|$ 在 $\mathcal{O}_\epsilon$ 上的一个上界; 由于 $D\boldsymbol{F}_{\boldsymbol{X}}$ 连续并且 $\mathcal{O}_\epsilon$ 紧, 因而存在这样的一个上界. 集合 $\mathcal{O}_\epsilon$ 是凸的, 也就是说, 如果 $\boldsymbol{Y},\boldsymbol{Z}\in\mathcal{O}_\epsilon$, 则连接 $\boldsymbol{Y}$ 和 $\boldsymbol{Z}$ 的线段整个地包含在 $\mathcal{O}_\epsilon$ 中. 这条线段可以写成 $\boldsymbol{Y}+s\boldsymbol{U}\in\mathcal{O}_\epsilon$, 其中 $\boldsymbol{U}=\boldsymbol{Z}-\boldsymbol{Y}$, $0\leqslant s\leqslant 1$. 令 $\psi(s)=\boldsymbol{F}(\boldsymbol{Y}+s\boldsymbol{U})$. 由链式法则可得,

$$\psi'(s)=D\boldsymbol{F}_{\boldsymbol{Y}+s\boldsymbol{U}}(\boldsymbol{U}).$$

从而,

$$\boldsymbol{F}(\boldsymbol{Z})-\boldsymbol{F}(\boldsymbol{Y})=\psi(1)-\psi(0)=\int_0^1\psi'(s)\mathrm{d}s=\int_0^1 D\boldsymbol{F}_{\boldsymbol{Y}+s\boldsymbol{U}}(\boldsymbol{U})\mathrm{d}s.$$

于是, 就有

$$|\boldsymbol{F}(\boldsymbol{Z})-\boldsymbol{F}(\boldsymbol{Y})|\leqslant\int_0^1 K|\boldsymbol{U}|\mathrm{d}s=K|\boldsymbol{Z}-\boldsymbol{Y}|.$$ ■

这个引理的证明同时还蕴涵：如果 $\mathcal{O}$ 是凸的, 并且对所有的 $\boldsymbol{X}\in\mathcal{O}$ 都有 $|D\boldsymbol{F}_{\boldsymbol{X}}|\leqslant K$, 则 K 就是 $\boldsymbol{F}|\mathcal{O}$ 的一个利普希茨常数.

假设 J 是一个包含零的开区间, 并且 $\boldsymbol{X}:J\to\mathcal{O}$ 满足

$$\boldsymbol{X}'(t)=\boldsymbol{F}(\boldsymbol{X}(t)),$$

其中 $\boldsymbol{X}(0)=\boldsymbol{X}_0$. 两边积分后, 有

$$\boldsymbol{X}(t)=\boldsymbol{X}_0+\int_0^t\boldsymbol{F}(\boldsymbol{X}(s))\mathrm{d}s.$$

这就是微分方程 $X' = F(X)$ 的积分形式. 反之, 如果 $X : J \to \mathcal{O}$ 满足上面的积分方程, 则有 $X(0) = X_0$, 并且微分之后可知, X 满足 $X' = F(X)$. 因而作为 $X : J \to \mathcal{O}$ 的方程, 积分形式和微分形式是等价的. 为了证明解的存在性, 我们将利用微分方程的积分形式.

现在我们来着手存在性的证明. 以下是我们的假设:

(1) $\mathcal{O}_\rho$ 是中心在 X_0 处、半径为 $\rho > 0$ 的闭球.

(2) F 在 $\mathcal{O}_\rho$ 上存在一个利普希茨常数 $K > 0$.

(3) 在 $\mathcal{O}_\rho$ 上有 $|F(X)| \leqslant M$.

(4) 选取正数 $a < \min\{\rho/M; 1/K\}$, 并记 $J = [-a, a]$.

我们首先定义从 J 到 $\mathcal{O}_\rho$ 的一串函数序列 $U_0, U_1, \cdots$. 然后我们将证明这列函数一致收敛并且极限函数满足我们的微分方程. 最后, 我们将证明没有其他这样的解. 用来证明 U_k 收敛性的引理如下:

分析引理 假设 $U_k : J \to \mathbb{R}^n$, $k = 0, 1, 2, \cdots$ 是定义在闭区间 J 上的一列连续函数, 并且它们满足: 任给 $\epsilon > 0$, 总存在 $N > 0$, 使得对所有的 $p, q > N$, 都有

$$\max_{t \in J} |U_p(t) - U_q(t)| < \epsilon.$$

则存在一个连续函数 $U : J \to \mathbb{R}^n$ 使得当 $k \to \infty$ 时有

$$\max_{t \in J} |U_k(t) - U(t)| \to 0.$$

并且, 对所有满足 $|t| \leqslant a$ 的 t, 都有

$$\lim_{k \to \infty} \int_0^t U_k(s) \mathrm{d}s = \int_0^t U(s) \mathrm{d}s. \qquad \blacksquare$$

函数列 U_k 的这种收敛性称为**一致收敛性**. 这个引理在一般的初等分析书中都有证明, 我们就不在此证明了. 参见 [38].

函数序列 U_k 是通过一种迭代手续递归地得到的, 这种迭代手续称为**毕卡迭代**. 在第 7 章中, 我们曾经给出了几个例子来解释毕卡迭代. 令

$$U_0(t) \equiv X_0.$$

对于 $t \in J$, 定义

$$U_1(t) = X_0 + \int_0^t F(U_0(s)) \mathrm{d}s = X_0 + tF(X_0).$$

由于 $|t| \leqslant a$ 以及 $|F(X_0)| \leqslant M$, 因而

$$|U_1(t) - X_0| = |t||F(X_0)| \leqslant aM \leqslant \rho,$$

因而对所有的 $t \in J$, 都有 $\boldsymbol{U}_1(t) \in \mathcal{O}_\rho$. 利用归纳法, 假设 $\boldsymbol{U}_k(t)$ 已经有定义, 并且对所有的 $t \in J$ 都有 $|\boldsymbol{U}_k(t) - \boldsymbol{X}_0| \leqslant \rho$. 然后, 令

$$\boldsymbol{U}_{k+1}(t) = \boldsymbol{X}_0 + \int_0^t \boldsymbol{F}(\boldsymbol{U}_k(s))\mathrm{d}s.$$

由于 $\boldsymbol{U}_k(s) \in \mathcal{O}_\rho$, 这样写是有意义的, 从而积分可以定义. 我们将证明 $|\boldsymbol{U}_{k+1}(t) - \boldsymbol{X}_0| \leqslant \rho$, 从而对所有的 $t \in J$, 都有 $\boldsymbol{U}_{k+1}(t) \in \mathcal{O}_\rho$; 这将意味着可以继续定义 $\boldsymbol{U}_{k+2}, \boldsymbol{U}_{k+3}$, 等等. 现在来证明 $|\boldsymbol{U}_{k+1}(t) - \boldsymbol{X}_0| \leqslant \rho$.

$$|\boldsymbol{U}_{k+1}(t) - \boldsymbol{X}_0| \leqslant \left|\int_0^t |\boldsymbol{F}(\boldsymbol{U}_k(s))|\mathrm{d}s\right| \leqslant \left|\int_0^t M\mathrm{d}s\right| \leqslant Ma < \rho.$$

下面证明, 存在一个常数 $L \geqslant 0$, 使得对所有的 $k \geqslant 0$, 都有

$$|\boldsymbol{U}_{k+1}(t) - \boldsymbol{U}_k(t)| \leqslant (aK)^k L.$$

令 L 为 $|\boldsymbol{U}_1(t) - \boldsymbol{U}_0(t)|$ 在 $-a \leqslant t \leqslant a$ 上的最大值. 根据上面的估计, 有 $L \leqslant aM$. 于是,

$$\begin{aligned}|\boldsymbol{U}_2(t) - \boldsymbol{U}_1(t)| &= \left|\int_0^t \boldsymbol{F}(\boldsymbol{U}_1(s)) - \boldsymbol{F}(\boldsymbol{U}_0(s))\mathrm{d}s\right| \\ &\leqslant \left|\int_0^t K|\boldsymbol{U}_1(s) - \boldsymbol{U}_0(s)|\mathrm{d}s\right| \\ &\leqslant aKL.\end{aligned}$$

现在归纳地假设对所有的 $k \geqslant 2$, 我们已经证明对所有的 $|t| \leqslant a$ 都有

$$|\boldsymbol{U}_k(t) - \boldsymbol{U}_{k-1}(t)| \leqslant (aK)^{k-1} L,$$

于是就有

$$\begin{aligned}|\boldsymbol{U}_{k+1}(t) - \boldsymbol{U}_k(t)| &\leqslant \left|\int_0^t |\boldsymbol{F}(\boldsymbol{U}_k(s)) - \boldsymbol{F}(\boldsymbol{U}_{k-1}(s))|\mathrm{d}s\right| \\ &\leqslant K\left|\int_0^t |\boldsymbol{U}_k(s) - \boldsymbol{U}_{k-1}(s)|\mathrm{d}s\right| \\ &\leqslant (aK)(aK)^{k-1} L \\ &= (aK)^k L.\end{aligned}$$

令 $\alpha = aK$, 根据定理的假设, 有 $\alpha < 1$. 任给 $\epsilon > 0$, 可以取 N 充分大, 使得对任何的 $r > s > N$, 都有

$$|\boldsymbol{U}_r(t) - \boldsymbol{U}_s(t)| \leqslant \sum_{k=N}^{\infty} |\boldsymbol{U}_{k+1}(t) - \boldsymbol{U}_k(t)| \leqslant \sum_{k=N}^{\infty} \alpha^k L \leqslant \epsilon,$$

最后一个不等式是因为, 对于 (收敛的) 几何级数, 总可以取 N 充分大, 使得之后的尾巴任意地小.

根据前面的分析引理, 这就说明了函数序列 $\boldsymbol{U}_0, \boldsymbol{U}_1, \cdots$ 一致收敛到一个连续函数 $\boldsymbol{X}: J \to \mathbb{R}^n$. 在等式

$$\boldsymbol{U}_{k+1}(t) = \boldsymbol{X}_0 + \int_0^t \boldsymbol{F}(\boldsymbol{U}_k(s))\mathrm{d}s$$

两边取极限可得

$$\begin{aligned}\boldsymbol{X}(t) &= \boldsymbol{X}_0 + \lim_{k\to\infty}\int_0^t \boldsymbol{F}(\boldsymbol{U}_k(s))\mathrm{d}s \\ &= \boldsymbol{X}_0 + \int_0^t \big(\lim_{k\to\infty}\boldsymbol{F}(\boldsymbol{U}_k(s))\big)\mathrm{d}s \\ &= \boldsymbol{X}_0 + \int_0^t \boldsymbol{F}(\boldsymbol{X}(s))\mathrm{d}s.\end{aligned}$$

其中的第二个等式也是根据分析引理得到的. 从而 $\boldsymbol{X}: J \to \mathcal{O}_\rho$ 满足微分方程的积分形式, 因而它也是微分方程的一个解. 特别地, 可以得出 $\boldsymbol{X}: J \to \mathcal{O}_\rho$ 是 C^1 的.

这就解决了定理的存在性部分. 现在我们来看唯一性.

假设 $\boldsymbol{X}, \boldsymbol{Y}: J \to \mathcal{O}$ 是微分方程的两个解, 并且满足 $\boldsymbol{X}(0) = \boldsymbol{Y}(0) = \boldsymbol{X}_0$, 这里的 J 与上面一样, 就是闭区间 $[-a, a]$. 我们将证明, 对所有的 $t \in J$, 都有 $\boldsymbol{X}(t) = \boldsymbol{Y}(t)$. 令

$$Q = \max_{t\in J}|\boldsymbol{X}(t) - \boldsymbol{Y}(t)|.$$

这个最大值可以在某一点 $t_1 \in J$ 达到. 于是,

$$\begin{aligned}Q = |\boldsymbol{X}(t_1) - \boldsymbol{Y}(t_1)| &= \left|\int_0^{t_1}(\boldsymbol{X}'(s) - \boldsymbol{Y}'(s))\mathrm{d}s\right| \\ &\leqslant \left|\int_0^{t_1}|\boldsymbol{F}(\boldsymbol{X}(s)) - \boldsymbol{F}(\boldsymbol{Y}(s))|\mathrm{d}s\right| \\ &\leqslant \left|\int_0^{t_1}K|\boldsymbol{X}(s) - \boldsymbol{Y}(s)|\mathrm{d}s\right| \\ &\leqslant aKQ.\end{aligned}$$

由于 $aK < 1$, 除非 $Q = 0$, 否则这是不可能的. 于是,

$$\boldsymbol{X}(t) \equiv \boldsymbol{Y}(t).$$

这就完成了定理的证明. ■

作为这个结果的总结, 我们已经证明, 任给中心在 $\boldsymbol{X}_0$、半径为 ρ 的闭球 $\mathcal{O}_\rho \subset \mathcal{O}$, 如果在 $\mathcal{O}_\rho$ 上满足:

(1) $|\boldsymbol{F}(\boldsymbol{X})| \leqslant M$,

(2) $\boldsymbol{F}$ 具有利普希茨常数 K,

(3) $0 < a < \min\{\rho/M, 1/K\}$,

则存在微分方程的唯一解 $\boldsymbol{X} : [-a, a] \to \mathcal{O}$ 使得 $\boldsymbol{X}(0) = \boldsymbol{X}_0$. 特别地, 如果 $\boldsymbol{F}$ 在 $\mathcal{O}$ 上为 C^1 的, 则这个结果成立.

下面依次给出几点说明. 首先, 如果 F 满足定理的假设, 则 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的两条解曲线不能相交. 这点是唯一性的直接推论, 但是值得从几何上加以强调. 假设 $\boldsymbol{X} : J \to \mathcal{O}$ 和 $\boldsymbol{Y} : J_1 \to \mathcal{O}$ 是 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的两个解, 并且 $\boldsymbol{X}(t_1) = \boldsymbol{Y}(t_2)$. 如果 $t_1 = t_2$, 则可由定理直接得出, 在 t_1 附近, $\boldsymbol{X}(t)$ 和 $\boldsymbol{Y}(t)$ 是相同的. 如果 $t_1 \neq t_2$, 令 $\boldsymbol{Y}_1(t) = \boldsymbol{Y}(t_2 - t_1 + t)$. 则 $\boldsymbol{Y}_1$ 也是该系统的一个解. 由于 $\boldsymbol{Y}_1(t_1) = \boldsymbol{Y}(t_2) = \boldsymbol{X}(t_1)$, 则根据定理的唯一性断言, 可知 $\boldsymbol{Y}_1$ 和 $\boldsymbol{X}$ 在 t_1 附近是相同的, 从而, 在 t_1 附近 $\boldsymbol{X}(t)$ 与 $\boldsymbol{Y}(t)$ 是重合的.

这里我们要强调一点, 如果 $\boldsymbol{Y}(t)$ 是一个解, 则对任意的常数 t_1, $\boldsymbol{Y}_1(t) = \boldsymbol{Y}(t + t_1)$ 也是一个解. 特别地, 如果 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的一条解曲线 $\boldsymbol{X} : J \to \mathcal{O}$ 对某个 t_1 和 $w > 0$ 满足 $\boldsymbol{X}(t_1) = \boldsymbol{X}(t_1 + w)$, 则这条解曲线实际上就是一个周期解, 也就是说, 对于所有的 t, 都有 $\boldsymbol{X}(t + w) = \boldsymbol{X}(t)$.

17.3 对初始条件的连续依赖性

为了使得解的存在唯一性定理在各种物理 (甚至是数学) 意义下都有意义, 还需要加上解 $\boldsymbol{X}(t)$ 对初值条件 $\boldsymbol{X}(0)$ 的连续依赖性质. 下面的定理给出了这个性质的准确陈述.

定理 设 $\mathcal{O} \subset \mathbb{R}^n$ 是开的, 并且 $F : \mathcal{O} \to \mathbb{R}^n$ 具有利普希茨常数 K. 设 $\boldsymbol{Y}(t)$ 和 $\boldsymbol{Z}(t)$ 是 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的两个解, 都在闭区间 $[t_0, t_1]$ 上有定义, 并且始终在 $\mathcal{O}$ 中. 则对所有的 $t \in [t_0, t_1]$, 都有

$$|\boldsymbol{Y}(t) - \boldsymbol{Z}(t)| \leqslant |\boldsymbol{Y}(t_0) - \boldsymbol{Z}(t_0)| \exp(K(t - t_0)). \qquad \blacksquare$$

这个结果表明, 如果两个解 $\boldsymbol{Y}(t)$ 和 $\boldsymbol{Z}(t)$ 出发时很接近, 则当 t 在 t_0 附近时, 它们一直很接近. 虽然这两个解可能分开, 但它们分开的速度不会超过指数增长. 特别地, 我们有下面的推论:

推论 (**对初始条件的连续依赖性**) 记 $\phi(t, \boldsymbol{X})$ 为系统 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$ 的流, 其中 $\boldsymbol{F}$ 为 C^1 的. 则 ϕ 为 $\boldsymbol{X}$ 的连续函数. $\blacksquare$

定理的证明依赖于下面的著名不等式, 我们先证明这个不等式.

Gronwall 不等式 设 $u:[0,\alpha]\to\mathbb{R}$ 为连续的, 并且非负. 假设存在 $C\geqslant 0$ 和 $K\geqslant 0$, 使得对所有的 $t\in[0,\alpha]$ 都有

$$u(t)\leqslant C+\int_0^t Ku(s)\mathrm{d}s.$$

则对所有的 $t\in[0,\alpha]$, 都有

$$u(t)\leqslant C\mathrm{e}^{Kt}.$$

证明 首先假设 $C>0$. 令

$$U(t)=C+\int_0^t Ku(s)\mathrm{d}s>0.$$

于是 $u(t)\leqslant U(t)$. 对 U 求微分可得

$$U'(t)=Ku(t).$$

于是,

$$\frac{U'(t)}{U(t)}=\frac{Ku(t)}{U(t)}\leqslant K.$$

从而就有

$$\frac{\mathrm{d}}{\mathrm{d}t}(\ln U(t))\leqslant K,$$

积分之后可得

$$\ln U(t)\leqslant \ln U(0)+Kt.$$

由于 $U(0)=C$, 取指数后就有

$$U(t)\leqslant C\mathrm{e}^{Kt},$$

从而就有

$$u(t)\leqslant C\mathrm{e}^{Kt}.$$

如果 $C=0$, 我们可以将上面的讨论用到一正数序列 c_i, 其中当 $i\to\infty$ 时, c_i 趋于 0. 这样就证明了 Gronwall 不等式. ■

证明 我们现在来证明定理. 定义

$$v(t)=|\boldsymbol{Y}(t)-\boldsymbol{Z}(t)|.$$

由于

$$\boldsymbol{Y}(t)-\boldsymbol{Z}(t)=\boldsymbol{Y}(t_0)-\boldsymbol{Z}(t_0)+\int_{t_0}^t \big(F(\boldsymbol{Y}(s))-F(\boldsymbol{Z}(s))\big)\mathrm{d}s,$$

我们有 (不妨设 $t \geqslant t_0$)

$$v(t) \leqslant v(t_0) + \int_{t_0}^{t} Kv(s)\mathrm{d}s.$$

令 $u(t) = v(t+t_0)$, 我们有

$$u(t) = v(t+t_0) \leqslant v(t_0) + \int_{t_0}^{t+t_0} Kv(s)\mathrm{d}s = v(t_0) + \int_0^t Ku(\tau)\mathrm{d}\tau,$$

对 $u(t)$ 运用 Gronwall 不等式可得 $v(t+t_0) \leqslant v(t_0)\exp(Kt)$, 也就是 $v(t) \leqslant v(t_0)\exp(K(t-t_0))$, 这正是定理的结论. ■

我们已经看到, 应用中产生的微分方程常常依赖于参数. 例如, 调和振子方程依赖于参数 b(阻尼系数) 和 k(弹性系数); 电路方程则依赖于电阻、电容以及感应系数等; 那么, 一个自然的问题就是: 这些方程的解如何依赖这些参数呢? 在前面的情形, 如果系统是以连续可微的方式依赖于这些参数, 则解就连续地依赖于这些参数. 利用一个特别的小技巧, 我们可以很容易地看出这一点. 假设系统

$$\boldsymbol{X}' = \boldsymbol{F}_a(\boldsymbol{X})$$

以 C^1 的方式依赖于参数 a. 我们来考虑一个“人为的”扩展的微分方程系统:

$$\begin{aligned} x_1' &= f_1(x_1, \cdots, x_n, a) \\ &\ \vdots \\ x_n' &= f_n(x_1, \cdots, x_n, a) \\ a' &= 0. \end{aligned}$$

现在这是一个 $n+1$ 个方程的自治系统. 尽管这个系统的扩展部分看起来很平凡, 但如果对它运用前面的解对初始条件的连续依赖性结果, 我们就可以得出原来系统的解对参数 a 也是连续依赖的.

定理 (**对参数的连续依赖性**) 令 $\boldsymbol{X}' = \boldsymbol{F}_a(\boldsymbol{X})$ 为一个微分方程系统, 其中 $\boldsymbol{F}_a$ 对 $\boldsymbol{X}$ 和 a 都连续可微. 则这个系统的流也连续依赖于 a, 就像它连续依赖于 $\boldsymbol{X}$ 一样. ■

17.4 延　伸　解

假设微分方程 $\boldsymbol{X}' = \boldsymbol{F}(\boldsymbol{X})$(其中 $\boldsymbol{F}$ 是 C^1 的) 有两个解 $\boldsymbol{Y}(t), \boldsymbol{Z}(t)$ 满足 $\boldsymbol{Y}(t_0) = \boldsymbol{Z}(t_0)$, 并且他们都定义在关于 t_0 的一个区间 J 上. 现在, 存在唯一性定理就保证, 对包含 t_0 的一个区间中的所有 t, 都有 $\boldsymbol{Y}(t) = \boldsymbol{Z}(t)$, 但这个区间可能比 J 要来得

小. 然而, 事实并非如此. 为了看出这一点, 假设 J^* 是使得 $\boldsymbol{Y}(t)=\boldsymbol{Z}(t)$ 的最大区间. 如果 $J^*\neq J$, 则存在 J^* 的一个端点 t_1 使得 $t_1\in J$. 根据连续性, 有 $\boldsymbol{Y}(t_1)=\boldsymbol{Z}(t_1)$. 现在, 定理的唯一性部分又保证, 事实上, $\boldsymbol{Y}(t)$ 和 $\boldsymbol{Z}(t)$ 在一个包含 t_1 的区间上都相同. 这与 J^* 是两个解相等的最大区间矛盾.

因此, 我们可以始终假设唯一解是定义在一个最大的时间区域上. 当然, 这并不能保证一个解 $\boldsymbol{X}(t)$ 对所有的时刻都有定义. 例如, 对任何常数 c, 函数 $x(t)=\tan(t-c)$ 都是微分方程

$$x'=1+x^2$$

的解. 但由于当 $t\to c\pm\pi/2$ 时, $x(t)\to\pm\infty$, 因而这样的解不能延伸到一个比

$$c-\frac{\pi}{2}<t<c+\frac{\pi}{2}$$

更大的区间上.

下面我们来研究当逼近一个解的定义域边界时, 这个解会出现什么情况. 我们只陈述右端点的结果; 左端点的情形是类似的.

定理 设 $\mathcal{O}\subset\mathbb{R}^n$ 是开的, 且 $F:\mathcal{O}\to\mathbb{R}^n$ 是 C^1 的. 设 $\boldsymbol{Y}(t)$ 是 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的一个解, 它定义在一个最大开区间 $J=(\alpha,\beta)\subset\mathbb{R}$ 上, 其中 $\beta<\infty$. 则任给一个紧集 $\mathcal{C}\subset\mathcal{O}$, 总存在某个 $t_0\in(\alpha,\beta)$, 使得 $\boldsymbol{Y}(t_0)\notin\mathcal{C}$.

这个定理告诉我们, 如果一个解 $\boldsymbol{Y}(t)$ 不能延伸到一个更大的时间区间, 则这个解要离开 $\mathcal{O}$ 中的任何紧集. 这意味着, 当 $t\to\beta$ 时, 要么 $\boldsymbol{Y}(t)$ 会在 $\mathcal{O}$ 的边界上聚集, 要么存在一个子序列 $t_i\to\beta$ 使得 $|\boldsymbol{Y}(t_i)|$ 趋于 ∞(或者两种可能都成立).

证明 假设对于所有的 $t\in(\alpha,\beta)$, 都有 $\boldsymbol{Y}(t)\in\mathcal{C}$. 由于 $\boldsymbol{F}$ 连续而且 $\mathcal{C}$ 是紧集, 因而存在 $M>0$ 使得对所有的 $\boldsymbol{X}\in\mathcal{C}$, 都有 $|\boldsymbol{F}(\boldsymbol{X})|\leqslant M$.

令 $\gamma\in(\alpha,\beta)$. 我们断言, $\boldsymbol{Y}$ 可以延伸为一个连续函数 $\boldsymbol{Y}:[\gamma,\beta]\to\mathcal{C}$. 为此, 我们只须证明 $\boldsymbol{Y}$ 在 J 上是一致连续的. 对于 $t_0,t_1\in J$, $t_0<t_1$, 有

$$|\boldsymbol{Y}(t_0)-\boldsymbol{Y}(t_1)=\left|\int_{t_0}^{t_1}\boldsymbol{Y}'(s)\mathrm{d}s\right|\leqslant\int_{t_0}^{t_1}|\boldsymbol{F}(\boldsymbol{Y}(s))|\mathrm{d}s\leqslant(t_1-t_0)M.$$

这就证明了 $\boldsymbol{Y}$ 在 J 上的一致连续性. 于是, 可以定义

$$\boldsymbol{Y}(\beta)=\lim_{t\to\beta}\boldsymbol{Y}(t).$$

接下来我们断言, 延伸后的曲线 $\boldsymbol{Y}:[\gamma,\beta]\to\mathbb{R}^n$ 在 β 处可微, 并且是微分方程的一个解. 我们有

$$\begin{aligned}\boldsymbol{Y}(\beta) &= \boldsymbol{Y}(\gamma)+\lim_{t\to\beta}\int_\gamma^t \boldsymbol{Y}'(s)\mathrm{d}s\\ &= \boldsymbol{Y}(\gamma)+\lim_{t\to\beta}\int_\gamma^t \boldsymbol{F}(\boldsymbol{Y}(s))\mathrm{d}s\\ &= \boldsymbol{Y}(\gamma)+\int_\gamma^\beta \boldsymbol{F}(\boldsymbol{Y}(s))\mathrm{d}s,\end{aligned}$$

这里我们用到了 $\boldsymbol{F}(\boldsymbol{Y}(s))$ 的一致连续性. 于是, 对位于 γ 和 β 之间的所有 t, 都有

$$\boldsymbol{Y}(t)=\boldsymbol{Y}(\gamma)+\int_\gamma^t \boldsymbol{F}(\boldsymbol{Y}(s))\mathrm{d}s.$$

因而, $\boldsymbol{Y}$ 在 β 处可微, 而且事实上, 有 $\boldsymbol{Y}'(\beta)=\boldsymbol{F}(\boldsymbol{Y}(\beta))$. 因此, $\boldsymbol{Y}$ 就是定义在 $[\gamma,\beta]$ 上的一个解. 因为对于某个 $\delta>\beta$, 总有一个定义在 $[\beta,\delta)$ 上的解, 因而可以将 $\boldsymbol{Y}$ 延伸到区间 (α,δ) 上. 于是, (α,β) 就不是一个解的最大定义域. 这就完成了定理的证明. ■

从这个定理可以立刻推出下面的重要事实:

推论 设 $\mathcal{C}$ 是开集 $\mathcal{O}\subset\mathbb{R}^n$ 中的一个紧子集, $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 是 C^1 的. 令 $\boldsymbol{Y}_0\in\mathcal{C}$, 并且假设任何一个形如 $\boldsymbol{Y}:[0,\beta]\to\mathcal{O},\boldsymbol{Y}(0)=\boldsymbol{Y}_0$ 的解曲线都整个地位于 $\mathcal{C}$ 中. 则存在一个解 $\boldsymbol{Y}:[0,\infty)\to\mathcal{O}$ 满足 $\boldsymbol{Y}(0)=\boldsymbol{Y}_0$, 并且对所有的 $t\geqslant 0$, 都有 $\boldsymbol{Y}(t)\in\mathcal{C}$, 于是, 这个解对所有的 (前向) 时间都有定义. ■

有了这些结果后, 对于解关于初始条件的连续性, 现在可以给出一个比在 17.3 节中讨论的结论要稍微强一些的定理. 在那里, 我们假设两个解在同一个区间上有定义. 在下面的定理中, 我们去掉了这个要求. 这个定理表明, 从邻近的点出发的解在同一个闭区间上有定义, 并且在这个区间上, 它们都彼此接近.

定理 设 $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 为 C^1 的. 假设 $\boldsymbol{Y}(t)$ 是 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的一个解, 在闭区间 $[t_0,t_1]$ 上有定义, 而且 $\boldsymbol{Y}(t_0)=\boldsymbol{Y}_0$. 则存在 $\boldsymbol{Y}_0$ 的一个邻域 $U\subset\mathbb{R}^n$ 以及常数 K 使得, 如果 $\boldsymbol{Z}_0\in U$, 则存在定义在 $[t_0,t_1]$ 上的唯一的解 $\boldsymbol{Z}(t)$ 满足 $\boldsymbol{Z}(t_0)=\boldsymbol{Z}_0$. 而且, 对所有的 $t\in[t_0,t_1]$, $\boldsymbol{Z}(t)$ 都满足

$$|\boldsymbol{Y}(t)-\boldsymbol{Z}(t)|\leqslant K|\boldsymbol{Y}_0-\boldsymbol{Z}_0|\exp(K(t-t_0)).$$ ■

我们需要下面的引理来证明这个定理.

引理 如果 $\boldsymbol{F}:\mathcal{O}\to\mathbb{R}^n$ 是局部利普希茨的, 而且 $\mathcal{C}\subset\mathcal{O}$ 是一个紧集, 则 $\boldsymbol{F}|\mathcal{C}$ 是利普希茨的.

证明 假设结论不成立, 则对每个 $K>0$, 无论多么大, 我们总可以在 $\mathcal{C}$ 中找到 $\boldsymbol{X}$ 和 $\boldsymbol{Y}$ 使得

$$|\boldsymbol{F}(\boldsymbol{X})-\boldsymbol{F}(\boldsymbol{Y})|>K|\boldsymbol{X}-\boldsymbol{Y}|.$$

特别地, 对于每个 $n=1,2,\cdots$, 可以找到 $\boldsymbol{X}_n,\boldsymbol{Y}_n$ 使得

$$|\boldsymbol{F}(\boldsymbol{X}_n)-\boldsymbol{F}(\boldsymbol{Y}_n)|>n|\boldsymbol{X}_n-\boldsymbol{Y}_n|.$$

由于 $\mathcal{C}$ 是紧的, 可以抽取 $\boldsymbol{X}_n$ 和 $\boldsymbol{Y}_n$ 的收敛子列. 重新标号后, 不妨假设 $\boldsymbol{X}_n\to\boldsymbol{X}^*,\boldsymbol{Y}_n\to\boldsymbol{Y}^*$, 其中 $\boldsymbol{X}^*,\boldsymbol{Y}^*\in\mathcal{C}$. 注意, 必定有 $\boldsymbol{X}^*=\boldsymbol{Y}^*$, 这是因为, 对所有的 n, 都有

$$|\boldsymbol{X}^*-\boldsymbol{Y}^*|=\lim_{n\to\infty}|\boldsymbol{X}_n-\boldsymbol{Y}_n|\leqslant\frac{1}{n}|\boldsymbol{F}(\boldsymbol{X}_n)-\boldsymbol{F}(\boldsymbol{Y}_n)|\leqslant\frac{2M}{n},$$

其中 M 是 $|\boldsymbol{F}(\boldsymbol{X})|$ 在 $\mathcal{C}$ 上的最大值. 存在 $\boldsymbol{X}^*$ 的邻域 $\mathcal{O}_0$ 使得 $\boldsymbol{F}|\mathcal{O}_0$ 具有利普希茨常数 K. 而且存在一个 n_0 使得当 $n\geqslant n_0$ 时, 有 $\boldsymbol{X}_n,\boldsymbol{Y}_n\in\mathcal{O}_0$. 从而, 对所有的 $n\geqslant n_0$, 有

$$|\boldsymbol{F}(\boldsymbol{X}_n)-\boldsymbol{F}(\boldsymbol{Y}_n)|\leqslant K|\boldsymbol{X}_n-\boldsymbol{Y}_n|,$$

当 $n\geqslant n_0$ 时, 这与前面的断言矛盾. 这就证明了引理. ■

证明 现在我们来证明定理. 根据 $[t_0,t_1]$ 的紧性, 存在 $\epsilon>0$, 使得只要对某个 $t\in[t_0,t_1]$, 有 $|\boldsymbol{X}-\boldsymbol{Y}(t)|\leqslant\epsilon$, 就可以推出 $\boldsymbol{X}\in\mathcal{O}$. 所有这样的点 $\boldsymbol{X}$ 构成了 $\mathcal{O}$ 的一个紧子集 $\mathcal{C}$. 我们在 17.2 节中已经看到, C^1 映射 F 是局部利普希茨的. 根据引理, 可知 $F|\mathcal{C}$ 具有一个利普希茨常数 K.

取 $\delta>0$ 充分小, 使得 $\delta\leqslant\epsilon$, 并且 $\delta\exp(K|t_1-t_0|)\leqslant\epsilon$. 我们断言, 如果 $|\boldsymbol{Z}_0-\boldsymbol{Y}_0|<\delta$, 则存在唯一一个过 $\boldsymbol{Z}_0$ 的解, 并且这个解在整个 $[t_0,t_1]$ 上都有定义. 首先, 由于 $|\boldsymbol{Z}_0-\boldsymbol{Y}(t_0)|<\epsilon$, 故 $\boldsymbol{Z}_0\in\mathcal{O}$, 因而存在一个过 $\boldsymbol{Z}_0$ 的解 $\boldsymbol{Z}(t)$, 它的最大定义区间为 $[t_0,\beta)$. 我们断言 $\beta>t_1$. 否则, 假设 $\beta\leqslant t_1$, 则由 Gronwall 不等式, 对所有的 $t\in[t_0,\beta)$, 都有

$$|\boldsymbol{Z}(t)-\boldsymbol{Y}(t)|\leqslant|\boldsymbol{Z}_0-\boldsymbol{Y}_0|\exp(K|t-t_0|)\leqslant\delta\exp(K|t-t_0|)\leqslant\epsilon.$$

从而 $\boldsymbol{Z}(t)$ 位于紧集 $\mathcal{C}$ 中. 根据上面的结果, $[t_0,\beta)$ 不可能是一个解的最大定义域. 因此, $\boldsymbol{Z}(t)$ 在 $[t_0,t_1]$ 上都有定义. $\boldsymbol{Z}(t)$ 的唯一性可由此立刻得出. 证毕. ■

17.5 非自治系统

在本节, 我们简要地考察一下非自治微分方程. 虽然本书的主要重点一直在自治系统上, 但是作为建立自治流可微性的一个技术工具, 非自治 (线性) 方程的理论是必需的.

设 $\mathcal{O} \subset \mathbb{R} \times \mathbb{R}^n$ 是一个开集, 又设 $\boldsymbol{F} : \mathcal{O} \to \mathbb{R}^n$ 是一个函数, 它关于 $\boldsymbol{X}$ 是 C^1 的, 但关于 t 可能仅仅只是连续的. 令 $(t_0, \boldsymbol{X}_0) \in \mathcal{O}$. 考虑非自治微分方程

$$\boldsymbol{X}'(t) = \boldsymbol{F}(t, \boldsymbol{X}), \quad \boldsymbol{X}(t_0) = \boldsymbol{X}_0.$$

照常, 该系统的一个解是指 $\mathbb{R}^n$ 中一条光滑曲线 $\boldsymbol{X}(t)$, 它定义在 t 的某个区间 J 上, 并且满足下面的性质:

(1) $t_0 \in J$, 且 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$;

(2) $(t, \boldsymbol{X}(t)) \in \mathcal{O}$, 并且对所有的 $t \in J$, 都有 $\boldsymbol{X}'(t) = \boldsymbol{F}(t, \boldsymbol{X}(t))$.

下面就是非自治方程的基本局部定理.

定理 设 $\mathcal{O} \subset \mathbb{R} \times \mathbb{R}^n$ 是开的, 又设 $\boldsymbol{F} : \mathcal{O} \to \mathbb{R}^n$ 是一个关于 $\boldsymbol{X}$ 为 C^1 且关于 t 连续的函数. 如果 $(t_0, \boldsymbol{X}_0) \in \mathcal{O}$, 则存在一个包含 t_0 的开区间 J, 以及 $\boldsymbol{X}' = \boldsymbol{F}(t, \boldsymbol{X})$ 定义在 J 上的唯一解, 满足 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0$. ■

它的证明与自治方程基本定理的证明 (17.2 节) 是相同的: 只需将额外的变量 t 插到适当的地方就可以了. 这个结果有如下的重要推论:

推论 设 $\boldsymbol{A}(t)$ 是 $n \times n$ 矩阵的一个连续族. 设 $(t_0, \boldsymbol{X}_0) \in J \times \mathbb{R}^n$. 则初值问题

$$\boldsymbol{X}' = \boldsymbol{A}(t)\boldsymbol{X}, \quad \boldsymbol{X}(t_0) = \boldsymbol{X}_0$$

在整个 J 上存在唯一的解. ■

对于它的一个证明, 见本章的习题 14.

我们称函数 $\boldsymbol{F}(t, \boldsymbol{X})$**关于 $\boldsymbol{X}$ 是利普希茨**的, 如果存在一个常数 $K \geqslant 0$, 使得对所有的 $(t, \boldsymbol{X}_1), (t, \boldsymbol{X}_2) \in \mathcal{O}$, 都有

$$|\boldsymbol{F}(t, \boldsymbol{X}_1) - \boldsymbol{F}(t, \boldsymbol{X}_2)| \leqslant K|\boldsymbol{X}_1 - \boldsymbol{X}_2|.$$

类似地, 可以定义关于 $\boldsymbol{X}$ 的局部利普希茨性.

与自治情形一样, 如果 $\boldsymbol{F}(t, \boldsymbol{X})$ 关于 $\boldsymbol{X}$ 是局部利普希茨的, 则非自治方程的解关于初始条件连续. 我们将这一事实的准确表述及其证明都留给读者.

解作为数据 $\boldsymbol{F}(t, \boldsymbol{X})$ 的函数的连续性是一种不同的连续性. 这就是说, 如果 $\boldsymbol{F} : \mathcal{O} \to \mathbb{R}^n$ 和 $\boldsymbol{G} : \mathcal{O} \to \mathbb{R}^n$ 关于 $\boldsymbol{X}$ 都是 C^1 的, 并且 $|\boldsymbol{F} - \boldsymbol{G}|$ 是一致的小, 我们期望 $\boldsymbol{X}' = \boldsymbol{F}(t, \boldsymbol{X})$ 和 $\boldsymbol{X}' = \boldsymbol{G}(t, \boldsymbol{X})$ 的初值相同的解也是接近的. 这是正确的, 事实上, 我们有下面更准确的结果:

定理 设 $\mathcal{O} \subset \mathbb{R} \times \mathbb{R}^n$ 是一个包含 $(0, \boldsymbol{X}_0)$ 的开集, 又设 $\boldsymbol{F}, \boldsymbol{G} : \mathcal{O} \to \mathbb{R}^n$ 关于 $\boldsymbol{X}$ 为 C^1 且关于 t 连续. 再假设对所有的 $(t, \boldsymbol{X}) \in \mathcal{O}$,

$$|\boldsymbol{F}(t, \boldsymbol{X}) - \boldsymbol{G}(t, \boldsymbol{X})| < \epsilon.$$

令 K 是 $\boldsymbol{F}(t,\boldsymbol{X})$ 关于 $\boldsymbol{X}$ 的一个利普希茨常数. 如果 $\boldsymbol{X}(t)$ 和 $\boldsymbol{Y}(t)$ 分别是方程 $\boldsymbol{X}'=\boldsymbol{F}(t,\boldsymbol{X})$ 和 $\boldsymbol{Y}'=\boldsymbol{G}(t,\boldsymbol{Y})$ 定义在某个区间 J 上的解, 并且 $\boldsymbol{X}(0)=\boldsymbol{X}_0=\boldsymbol{Y}(0)$, 则对所有的 $t\in J$, 都有

$$|\boldsymbol{X}(t)-\boldsymbol{Y}(t)|\leqslant\frac{\epsilon}{K}\big(\exp(K|t|)-1\big).$$

证明 对 $t\in J$, 有

$$\begin{aligned}\boldsymbol{X}(t)-\boldsymbol{Y}(t)&=\int_0^t(\boldsymbol{X}'(s)-\boldsymbol{Y}'(s))\mathrm{d}s\\&=\int_0^t(\boldsymbol{F}(s,\boldsymbol{X}(s))-\boldsymbol{G}(s,\boldsymbol{Y}(s)))\mathrm{d}s.\end{aligned}$$

于是, 当 $t\geqslant 0$ 时, 有

$$\begin{aligned}|\boldsymbol{X}(t)-\boldsymbol{Y}(t)|&\leqslant\int_0^t|\boldsymbol{F}(s,\boldsymbol{X}(s))-\boldsymbol{F}(s,\boldsymbol{Y}(s))|\mathrm{d}s\\&\quad+\int_0^t|\boldsymbol{F}(s,\boldsymbol{Y}(s))-\boldsymbol{G}(s,\boldsymbol{Y}(s))|\mathrm{d}s\\&\leqslant\int_0^t K|\boldsymbol{X}(s)-\boldsymbol{Y}(s)|\mathrm{d}s+\int_0^t\epsilon\mathrm{d}s.\end{aligned}$$

令 $u(t)=|\boldsymbol{X}(t)-\boldsymbol{Y}(t)|$. 则

$$u(t)\leqslant K\int_0^t\left(u(s)+\frac{\epsilon}{K}\right)\mathrm{d}s,$$

从而,

$$u(t)+\frac{\epsilon}{K}\leqslant\frac{\epsilon}{K}+K\int_0^t\left(u(s)+\frac{\epsilon}{K}\right)\mathrm{d}s.$$

由 Gronwall 不等式可得

$$u(t)+\frac{\epsilon}{K}\leqslant\frac{\epsilon}{K}\exp(K|t|),$$

这就在 $t\geqslant 0$ 的情形证明了定理. $t<0$ 的情形可以类似地处理. ∎

17.6 流的可微性

现在我们再回到自治微分方程 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 的情形, 这里假设 F 是 C^1 的. 我们的目的是证明该方程所确定的流 $\phi(t,\boldsymbol{X})=\phi_t(\boldsymbol{X})$ 是两个变量的 C^1 函数, 同时我们还将确定 $\partial\phi/\partial\boldsymbol{X}$. 当然, 我们知道 ϕ 关于变量 t 是连续可微的, 因而只需证明关于 $\boldsymbol{X}$ 的可微性.

为此, 设 $\boldsymbol{X}(t)$ 是该系统的一个特定的解, 在关于 0 的一个闭区间 J 上对 t 有定义. 设 $\boldsymbol{X}(0)=\boldsymbol{X}_0$. 对每个 $t\in J$, 令

$$\boldsymbol{A}(t)=D\boldsymbol{F}_{\boldsymbol{X}(t)}.$$

即, $\boldsymbol{A}(t)$ 代表 $\boldsymbol{F}$ 在点 $\boldsymbol{X}(t)$ 处的雅可比矩阵. 由于 $\boldsymbol{F}$ 为 C^1 的, 故 $\boldsymbol{A}(t)$ 连续. 我们定义非自治线性方程

$$\boldsymbol{U}'=\boldsymbol{A}(t)\boldsymbol{U}.$$

这个方程称为沿解 $\boldsymbol{X}(t)$ 的**变分方程**. 由上一节可知, 对于每个初始条件 $\boldsymbol{U}(0)=\boldsymbol{U}_0$, 变分方程都有一个定义在整个 J 上的解. 而且, 与自治情形一样, 变分方程的解满足线性叠加原理.

变分方程的意义在于, 如果 $\boldsymbol{U}_0$ 很小, 则函数

$$t\to\boldsymbol{X}(t)+\boldsymbol{U}(t)$$

很好地近似于原来的自治方程满足初值 $\boldsymbol{X}(0)=\boldsymbol{X}_0+\boldsymbol{U}_0$ 的解 $\boldsymbol{X}(t)$.

为了准确起见, 假设 $\boldsymbol{U}(t,\boldsymbol{\xi})$ 是变分方程满足 $\boldsymbol{U}(0,\boldsymbol{\xi})=\boldsymbol{\xi}$ 的解, 其中 $\boldsymbol{\xi}\in\mathbb{R}^n$. 如果 $\boldsymbol{\xi}$ 和 $\boldsymbol{X}_0+\boldsymbol{\xi}$ 都属于 $\mathcal{O}$, 令 $\boldsymbol{Y}(t,\boldsymbol{\xi})$ 是自治方程 $\boldsymbol{X}'=\boldsymbol{F}(\boldsymbol{X})$ 满足 $\boldsymbol{Y}(0)=\boldsymbol{X}_0+\boldsymbol{\xi}$ 的解.

命题 令 J 是一个包含 0 的闭区间, 并且 $\boldsymbol{X}(t)$ 在上面有定义. 则

$$\lim_{\boldsymbol{\xi}\to\boldsymbol{0}}\frac{|\boldsymbol{Y}(t,\boldsymbol{\xi})-\boldsymbol{X}(t)-\boldsymbol{U}(t,\boldsymbol{\xi})|}{|\boldsymbol{\xi}|}$$

对于 $t\in J$ 一致地趋于 0. ■

这意味着, 对每个 $\epsilon>0$, 都存在 $\delta>0$, 使得只要 $|\boldsymbol{\xi}|\leqslant\delta$, 则对所有的 $t\in J$, 都有

$$|\boldsymbol{Y}(t,\boldsymbol{\xi})-(\boldsymbol{X}(t)+\boldsymbol{U}(t,\boldsymbol{\xi}))|\leqslant\epsilon|\boldsymbol{\xi}|.$$

因而当 $\boldsymbol{\xi}\to\boldsymbol{0}$ 时, 曲线 $t\to\boldsymbol{X}(t)+\boldsymbol{U}(t,\boldsymbol{\xi})$ 是 $\boldsymbol{Y}(t,\boldsymbol{\xi})$ 的一个越来越好的近似. 在许多应用中, $\boldsymbol{X}(t)+\boldsymbol{U}(t,\boldsymbol{\xi})$ 被用来代替 $\boldsymbol{Y}(t,\boldsymbol{\xi})$; 由于 $\boldsymbol{U}(t,\boldsymbol{\xi})$ 关于 $\boldsymbol{\xi}$ 线性, 这种代替将会带来方便.

我们稍后再证明这个命题, 在此之前, 我们先利用这个结果来证明下面的定理:

定理 **(流的可微性)** 自治系统 $\boldsymbol{X}'=F(\boldsymbol{X})$ 的流 $\phi(t,\boldsymbol{X})$ 是一个 C^1 函数; 即, $\partial\phi/\partial t$ 和 $\partial\phi/\partial\boldsymbol{X}$ 都存在, 而且都对 t 和 $\boldsymbol{X}$ 连续.

证明 当然, $\partial\phi(t,\boldsymbol{X})/\partial t$ 就是 $F(\phi(t,\boldsymbol{X}))$, 因而是连续的. 为了计算 $\partial\phi(t,\boldsymbol{X})/\partial\boldsymbol{X}$, 我们有, 对很小的 $\boldsymbol{\xi}$,

$$\phi(t,\boldsymbol{X}_0+\boldsymbol{\xi})-\phi(t,\boldsymbol{X}_0)=\boldsymbol{Y}(t,\boldsymbol{\xi})-\boldsymbol{X}(t).$$

命题现在蕴含 $\partial\phi(t,\boldsymbol{X}_0)/\partial\boldsymbol{X}$ 就是线性映射 $\boldsymbol{\xi}\to\boldsymbol{U}(t,\boldsymbol{\xi})$. 于是, $\partial\phi/\partial\boldsymbol{X}$ 的连续性可以从变分方程的解对初始条件的连续性以及关于数据函数 $\boldsymbol{A}(t)=D\boldsymbol{F}_{\boldsymbol{X}(t)}$ 的连续性推出. ■

现在将流记为 $\phi_t(\boldsymbol{X})$, 注意, 对于每个 t, 映射 ϕ_t 在 $\boldsymbol{X}\in\mathcal{O}$ 处的导算子 $D\phi_t(\boldsymbol{X})$ 就是 $\partial\phi(t,\boldsymbol{X})/\partial\boldsymbol{X}$. 我们将它称为这个流的**空间导数**, 相应地, 称 $\partial\phi(t,\boldsymbol{X})/\partial t$ 为**时间导数**.

上一定理的证明实际上说明了 $D\phi_t(\boldsymbol{X})$ 是 $\mathbb{R}^n$ 的线性映射空间上的一个初值问题的解: 对每个 $\boldsymbol{X}_0\in\mathcal{O}$, 流的空间导数满足微分方程

$$\frac{\mathrm{d}}{\mathrm{d}t}(D\phi_t(\boldsymbol{X}_0))=DF_{\phi_t(\boldsymbol{X}_0)}D\phi_t(\boldsymbol{X}_0),$$

以及初始条件 $D\phi_0(\boldsymbol{X}_0)=\boldsymbol{I}$. 这里我们将 $\boldsymbol{X}_0$ 看成一个参数.

一个重要的特殊情形就是一个平衡解 $\overline{\boldsymbol{X}}$ 的情形, 此时 $\phi_t(\overline{\boldsymbol{X}})\equiv\overline{\boldsymbol{X}}$. 令 $D\boldsymbol{F}_{\overline{\boldsymbol{X}}}=\boldsymbol{A}$, 我们得到微分方程

$$\frac{\mathrm{d}}{\mathrm{d}t}(D\phi_t(\overline{\boldsymbol{X}}))=\boldsymbol{A}D\phi_t(\overline{\boldsymbol{X}}),$$

以及 $D\phi_0(\overline{\boldsymbol{X}})=\boldsymbol{I}$. 该方程的解为

$$D\phi_t(\overline{\boldsymbol{X}})=\exp(t\boldsymbol{A}).$$

这意味着, 在平衡点的一个邻域中, 流近似线性.

现在我们来证明命题. $\boldsymbol{X}(t),\boldsymbol{Y}(t,\boldsymbol{\xi})$ 和 $\boldsymbol{U}(t,\boldsymbol{\xi})$ 满足的积分方程分别是

$$\begin{aligned}\boldsymbol{X}(t)&=\boldsymbol{X}_0+\int_0^t\boldsymbol{F}(\boldsymbol{X}(s))\mathrm{d}s,\\ \boldsymbol{Y}(t,\boldsymbol{\xi})&=\boldsymbol{X}_0+\boldsymbol{\xi}+\int_0^t\boldsymbol{F}(\boldsymbol{Y}(s,\boldsymbol{\xi}))\mathrm{d}s,\\ \boldsymbol{U}(t,\boldsymbol{\xi})&=\boldsymbol{\xi}+\int_0^t D\boldsymbol{F}_{\boldsymbol{X}(s)}(\boldsymbol{U}(s,\boldsymbol{\xi}))\mathrm{d}s.\end{aligned}$$

从这些等式可得, 对所有的 $t\geqslant 0$, 都有

$$|\boldsymbol{Y}(t,\boldsymbol{\xi})-\boldsymbol{X}(t)-\boldsymbol{U}(t,\boldsymbol{\xi})|\leqslant\int_0^t|\boldsymbol{F}(\boldsymbol{Y}(s,\boldsymbol{\xi}))-\boldsymbol{F}(\boldsymbol{X}(s))-D\boldsymbol{F}_{\boldsymbol{X}(s)}(\boldsymbol{U}(s,\boldsymbol{\xi}))|\mathrm{d}s.$$

$\boldsymbol{F}$ 在一点 $\boldsymbol{Z}$ 的泰勒逼近为

$$\boldsymbol{F}(\boldsymbol{Y})=\boldsymbol{F}(\boldsymbol{Z})+D\boldsymbol{F}_{\boldsymbol{Z}}(\boldsymbol{Y}-\boldsymbol{Z})+R(\boldsymbol{Z},\boldsymbol{Y}-\boldsymbol{Z}),$$

其中的余项满足: 对于一个给定紧集上的 $\boldsymbol{Z}$ 一致地有

$$\lim_{\boldsymbol{Y}\to\boldsymbol{Z}}\frac{R(\boldsymbol{Z},\boldsymbol{Y}-\boldsymbol{Z})}{|\boldsymbol{Y}-\boldsymbol{Z}|}=0.$$

对 $\boldsymbol{Y}=\boldsymbol{Y}(s,\boldsymbol{\xi}), \boldsymbol{Z}=\boldsymbol{X}(s)$ 运用上面的公式, 根据 $D\boldsymbol{F}_{\boldsymbol{X}(s)}$ 的线性性, 可得

$$|\boldsymbol{Y}(t,\boldsymbol{\xi})-\boldsymbol{X}(t)-\boldsymbol{U}(t,\boldsymbol{\xi})| \leqslant \int_0^t |D\boldsymbol{F}_{\boldsymbol{X}(s)}(\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s)-\boldsymbol{U}(s,\boldsymbol{\xi}))|\mathrm{d}s + \int_0^t |R(\boldsymbol{X}(s),\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s))|\mathrm{d}s.$$

将上面的表达式左边记为 $g(t)$, 并且令

$$N=\max\{|D\boldsymbol{F}_{\boldsymbol{X}(s)}| \mid s\in J\}.$$

于是就有

$$g(t)\leqslant N\int_0^t g(s)\mathrm{d}s+\int_0^t |R(\boldsymbol{X}(s),\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s))|\mathrm{d}s.$$

固定 $\epsilon>0$ 并选取 $\delta_0>0$ 充分小, 使得只要 $|\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s)|\leqslant\delta_0$, $s\in J$, 就有

$$|R(\boldsymbol{X}(s),\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s))|\leqslant\epsilon|\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s)|.$$

从 17.3 节可知, 存在常数 $K\geqslant 0$ 和 $\delta_1>0$, 使得只要 $|\boldsymbol{\xi}|\leqslant\delta_1, s\in J$ 就有

$$|\boldsymbol{Y}(s,\boldsymbol{\xi})-\boldsymbol{X}(s)|\leqslant|\boldsymbol{\xi}|\mathrm{e}^{Ks}\leqslant\delta_0.$$

现在假设 $|\boldsymbol{\xi}|\leqslant\delta_1$. 从前面的方程可知, 对 $t\in J$ 有

$$g(t)\leqslant N\int_0^t g(s)\mathrm{d}s+\int_0^t \epsilon|\boldsymbol{\xi}|\mathrm{e}^{Ks}\mathrm{d}s,$$

于是

$$g(t)\leqslant N\int_0^t g(s)\mathrm{d}s+C\epsilon|\boldsymbol{\xi}|,$$

其中 C 是某个常数, 它只依赖于 K 和区间 J 的长度. 根据 Gronwall 不等式可得, 只要 $t\in J$ 且 $|\boldsymbol{\xi}|\leqslant\delta_1$ 就有

$$g(t)\leqslant C\epsilon\mathrm{e}^{Nt}|\boldsymbol{\xi}|.$$

(注意, δ_1 依赖于 ϵ.) 由于 ϵ 是任意的正数, 这就证明了 $g(t)/|\boldsymbol{\xi}|$ 在 $t\in J$ 中且 $t\geqslant 0$ 时一致趋于 0. 对于 $t\leqslant 0$ 的情形可以类似证明. 这也就证明了命题.

习 题

1. 对下面的每个初值问题, 写出毕卡迭代序列的前几项. 在可能的情形, 利用各种方法找出显式解. 讨论解的定义域.

(a) $x' = x - 2, \quad x(0) = 1$;

(b) $x' = x^{4/3}, \quad x(0) = 0$;

(c) $x' = x^{4/3}, \quad x(0) = 1$;

(d) $x' = \cos x, \quad x(0) = 0$;

(e) $x' = 1/(2x), \quad x(1) = 1$.

2. 设 $\boldsymbol{A}$ 是一个 $n \times n$ 矩阵. 证明用毕卡迭代法求解 $\boldsymbol{X}' = \boldsymbol{A}\boldsymbol{X}$, $\boldsymbol{X}(0) = \boldsymbol{X}_0$ 给出的解就是 $\exp(t\boldsymbol{A})\boldsymbol{X}_0$.

3. 对二阶初值问题

$$x'' = -x; \quad x(0) = 1, \quad x'(0) = 0$$

对应的一阶系统运用毕卡迭代, 进而导出 $\cos t$ 的泰勒级数.

4. 对于下面的每个函数, 找出指定区域上的一个利普希茨常数, 或者证明不存在利普希茨常数.

(a) $f(x) = |x|, \quad -\infty < x < \infty$;

(b) $f(x) = x^{1/3}, \quad -1 \leqslant x \leqslant 1$;

(c) $f(x) = 1/x, \quad 1 \leqslant x < \infty$;

(d) $f(x, y) = (x + 2y, -y), \quad (x, y) \in \mathbb{R}^2$;

(e) $f(x, y) = \dfrac{xy}{1 + x^2 + y^2}, \quad x^2 + y^2 \leqslant 4$.

5. 考虑微分方程

$$x' = x^{1/3}.$$

它有多少个不同的解满足 $x(0) = 0$?

6. 关于微分方程 $x' = x/t$ 的解能够说些什么?

7. 定义 $f : \mathbb{R} \to \mathbb{R}$ 如下：如果 $x \leqslant 1$, 则 $f(x) = 1$; 如果 $x > 1$, 则 $f(x) = 2$. 关于 $x' = f(x)$ 满足 $x(0) = 1$ 的解能够说些什么? 注意, 这个微分方程的右端不连续. 如果改成：当 $x > 1$ 时, $f(x) = 0$, 会发生什么?

8. 设 $\boldsymbol{A}(t)$ 是 $n \times n$ 矩阵的一个连续族, 并且 $\boldsymbol{P}(t)$ 是初值问题 $\boldsymbol{P}' = \boldsymbol{A}(t)\boldsymbol{P}$, $\boldsymbol{P}(0) = \boldsymbol{P}_0$ 的矩阵解. 证明：

$$\det \boldsymbol{P}(t) = (\det \boldsymbol{P}_0) \exp\left(\int_0^t \mathrm{Tr}\boldsymbol{A}(s)\mathrm{d}s\right).$$

9. 假设 $\boldsymbol{F}$ 是一个梯度向量场. 证明 $|D\boldsymbol{F}_{\boldsymbol{X}}|$ 就是 $D\boldsymbol{F}_{\boldsymbol{X}}$ 的最大特征值 (在绝对值意义下). (提示：$D\boldsymbol{F}_{\boldsymbol{X}}$ 是一个对称矩阵.)

10. 证明：二阶的两点边值问题

$$x'' = -x; \quad x(0) = 0, \quad x(\pi) = 1$$

无解.

11. 如果将上题中的微分方程替换成 $x''=-kx(k>0)$, 会发生什么情况?
12. 证明下面的一般事实 (参见 17.3 节): 如果 $C\geqslant 0$, $u,v:[0,\beta]\to\mathbb{R}$ 都连续且非负, 并且对所有的 $t\in[0,\beta]$ 都有

$$u(t)\leqslant C+\int_0^t u(s)v(s)\mathrm{d}s,$$

则, $u(t)\leqslant C\mathrm{e}^{V(t)}$, 其中

$$V(t)=\int_0^t v(s)\mathrm{d}s.$$

13. 假设 $\mathcal{C}\subset\mathbb{R}^n$ 紧的, 且 $f:\mathcal{C}\to\mathbb{R}$ 连续. 证明 f 在 $\mathcal{C}$ 上有界, 并且 f 在 $\mathcal{C}$ 中的某点达到它的最大值.
14. 设 $\boldsymbol{A}(t)$ 是 $n\times n$ 矩阵的一个连续族. 令 $(t_0,\boldsymbol{X}_0)\in J\times\mathbb{R}^n$. 则初值问题

$$\boldsymbol{X}'=\boldsymbol{A}(t)\boldsymbol{X},\quad \boldsymbol{X}(t_0)=\boldsymbol{X}_0$$

在整个 J 上有唯一解.
15. 用一篇不超过 50 页的论文描述系统 $\boldsymbol{X}'=0$ 所有解的行为, 其中 $\boldsymbol{X}\in\mathbb{R}^n$. 噢, 没错. 这是数学系送给你的另一件免费的, 也是最后的一件礼物.

参考文献

1. Abraham, R. and Marsden, J., *Foundations of Mechanics*. Reading, MA: Benjamin-Cummings, 1978.

2. Abraham, R. and Shaw, C., *Dynamics : The Geometry of Behavior*. Redwood City, CA: Addison-Wesley, 1992.

3. Alligood, K., Sauer, T. and Yorke, J., *Chaos: An Introduction to Dynamical Systems*. New York: Springer-Verlag, 1997.

4. Afraimovich, V. S. and Shil'nikov, L.P., Strange attractors and quasiattractors. In *Nonlinear Dynamics and Turbulence*. Boston: Pitman, (1983), 1.

5. Arnold, V. I., *Ordinary Differential Equations*. Cambridge: MIT Press, 1973.

6. Arnold, V. I., *Mathematical Methods of Classical Mechanics*. New York: Springer-Verlag, 1978.

7. Arrowsmith, D. and Place, C., *An Introduction to Dynamical Systems*. Cambridge: Cambridge University Press, 1990.

8. Banks, J. *et al.* On Devaney's definition of chaos. *Amer. Math. Monthly*. **99** (1992), 332.

9. Birman, J.S. and Williams, R.F., Knotted periodic orbits in dynamical systems I: Lorenz's equations. *Topology*. **22**(1983), 47.

10. Blanchard, P., Devaney, R. L. and Hall, G. R., *Differential Equations*. Pacific Grove, CA: Brooks-Cole, 2002.

11. Chua, L., Komuro, M. and Matsumoto, T., The double scroll family. *IEEE Trans. on Circuits and Systems*. **33**(1986), 1073.

12. Coddington, E. and Levinsov, N., *Theory of Ordinary Equations*. New York: McGraw-Hill, 1955.

13. Devaney, R. L., *Introduction to Chaotic Dynamical Systems*. Boulder, CO: Westview Press, 1989.

14. Devaney, Kℓ. Math texts and digestion. *J. Obesity*. **23**(2002), 1.8.

15. Edelstein-Keshet, L., *Mathematical Models in Biology*. New York: McGraw-Hill, 1987.

16. Ermentrout, G. B. and Kopell, N., Oscillatar death in systems of coupled neural oscillators. *SIAM J. Appl. Math*. **50**(1990), 125.

17. Field, R. and Burger, M., eds. *Oscillations and Traveling Waves in Chemical Systems*. New York: Wiley, 1985.

18. Fitzhugh, R., Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* **1**(1961), 445.

19. Golubitsky, M., Josić, K. and Kaper, T., An unfolding theory approach to bursting in fast-slow systems. In *Global Theory of Dynamical Systems*. Bristol, UK: Institute of Physics, 2001, 277.

20. Guckenheimer, J. and Williams, R. F., Structural stability of Lorenz attractors. *Publ. Math. IHES.* **50**(1979), 59.

21. Guckenheimer, J. and Holmes, P., *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York: Springer-Verlag, 1983.

22. Gutzwiller, M., The anisotropic Kepler problem in two dimensions. *J. Math. Phys.* **14** (1973), 139.

23. Hodgkin, A. L. and Huxley, A. F., A quantitative description of membrane current and its application to conduction and excitation in nerves. *J. Physiol.* **117**(1952), 500.

24. Katok, A. and Hasselblatt, B., *Introduction to the Modern Theory of Dynamical Systems*. Cambridge, UK: Cambridge University Press, 1995.

25. Khibnik, A., Roose, D. and Chua, L., On periodic orbits and homoclinic bifurcations in Chua's circuit with a smooth nonlinearity. *Int. J. Bifurcation and Chaos.* **3** (1993), 363.

26. Kraft, R., Chaos, Cantor sets, and hyperbolicity for the logistic maps. *Amer. Math Montnly.* **106**(1999), 400.

27. Lengyel, I., Rabai, G. and Epstein, I., Experimental and modeling study of oscillations in the chlorine dioxide-iodine-malonic acid reaction. *J. Amer. Chem. Soc.* **112**(1990), 9104.

28. Liapunov, A. M., *The General Problem of Stability of Motion*. London: Taylor & Francis, 1992.

29. Lorenz, E., Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**(1963), 130.

30. Marsden, J. E. and McCracken, M., *The Hopf Bifurcation and Its Applications*. New York: Springer-Verlag, 1976.

31. May, R. M., *Theoretical Ecology : Principles and Applications*. Oxford: Blackwell, 1981.

32. McGehee, R., Triple collision in the collinear three body problem. *Inventiones Math.* **27**(1974), 191.

33. Moeckel, R., Chaotic dynamics near triple collision. *Arch. Rational Mech. Anal.* **107** (1989), 37.

34. Murray, J. D., *Mathematical Biology*. Berlin: Springer-Verlag, 1993.

35. Nagumo, J. S., Arimoto, S. and Yoshizawa, S., An active pulse transmission line stimulating nerve axon. *Proc. IRE.* **50**(1962), 2061.

36. Robinson, C., *Dynamical Systems : Stability, Symbolic Dynamics, and Chaos*. Boca Raton, FL: CRC Press, 1995.

37. Rössler, O. E, An equation for continuous chaos. *Phys. Lett.A* **57**(1976), 397.

38. Rudin, W., *Principles of Mathematical Analysis*. New York: McGraw-Hill, 1976.

39. Schneider, G. and Wayne, C. E., Kawahara dynamics in dispersive media. *Phys. D* **152**(2001), 384.

40. Shil'nikov, L. P., A case of the existence of a countable set of periodic motions. Sov.*Math.Dokl.*6(1965), 163.

41. Shil'nikov, L.P., Chua's circuit: Rigorous results and future problems. *Int. J. Bifurcation and Chaos*.4(1994), 489.

42. Siegel, C. and Moser, J., *Lectures on Celestial Mechanics*. Berlin: Springer-Verlag, 1971.

43. Smale, S., Diffeomorphisms with many periodic points. In *Differential and Combinatorial Topology*. Princeton, NJ: Princeton University Press, 1965, 63.

44. Sparrow, C., *The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors*. New York: Springer-Verlag, 1982.

45. Strogatz, S., *Nonlinear Dynamics and Chaos*. Reading, MA: Addison-Wesley, 1994.

46. Tucker, W., The Lorenz attractor exists. *C. R. Acad. Sci. Paris Sér. I Math.* **328**(1999), 1197.

47. Winfree A. T., The prehistory of the Belousov-Zhabotinsky reaction. *J. Chem. Educ.* **61**(1984), 661.

索引

M

N

O

P

Q

S

其他